1993

General Tests of Latent Variable Models and Mean-Variance Spanning

Wayne E. Ferson

Stephen R. Foerster

Donald B. Keim
University of Pennsylvania

Recommended Citation

Ferson, W. E., Foerster, S. R., & Keim, D. B. (1993). General Tests of Latent Variable Models and Mean-Variance Spanning. *The Journal of Finance, 48* (1), 131-156. http://dx.doi.org/10.1111/j.1540-6261.1993.tb04704.x

General Tests of Latent Variable Models and Mean-Variance Spanning

Abstract
The methods of Gibbons and Ferson (1985) are extended, relaxing the assumption that expected returns are linear functions of predetermined instruments. A model of conditional mean-variance spanning generalizes Huberman and Kandel (1987). The empirical results indicate that more than a single risk premium is needed to model expected stock and bond returns, but the number of common factors in the expected returns is small. However, when size-based common stock portfolios proxy for the risk factors, we reject the hypothesis that four of them describe the conditional expected returns of the other assets.

Disciplines
Finance | Finance and Financial Management

GENERAL TESTS OF LATENT VARIABLE MODELS AND MEAN VARIANCE SPANNING

by

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

10-92

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

current version: June 5, 1992

forthcoming: Journal of Finance

General Tests of Latent Variable Models and Mean Variance Spanning

WAYNE E. FERSON, STEPHEN R. FOERSTER and DONALD B. KEIM

ABSTRACT

The methods of Gibbons and Ferson (1985) are extended, relaxing the assumption that expected returns are linear functions of predetermined instruments. A model of conditional mean variance spanning generalizes Huberman and Kandel (1987). The empirical results indicate that more than a single risk premium is needed to model expected stock and bond returns, but the number of common factors in the expected returns is small. However, when size-based common stock portfolios proxy for the risk factors, we reject the hypothesis that four of them describe the conditional expected returns of the other assets.

Graduate School of Business, University of Chicago, 1101 East 58-th Street, Chicago, IL 60637 (312-702-7039), School of Business Administration, University of Western Ontario and the Wharton School, University of Pennsylvania, respectively. Proofs and reprint order forms to Ferson.
Formerly circulated under the title: "Tests of Asset Pricing Models with Changing Expectations." We thank Phillip Braun, Mark Grinblatt, Robert Korajczyk, Andrew Lo, Craig MacKinlay, two anonymous referees 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. The first author acknowledges financial support from the Center for Research in Security Prices. The second author acknowledges financial support from the Plan for Excellence.

This paper makes two contributions to the burgeoning literature on tests of asset pricing models with changing expected returns. First, we extend and generalize the latent variables methodology of Gibbons and Ferson (1985, henceforth GF). Latent variable models assume that expected returns vary over time as functions of a small number of risk premiums, which are common across assets. The expected risk premiums are treated as unobserved latent variables. Numerous studies have applied such models to study the expected returns of stocks, bonds, foreign exchange and other assets.[1] Our general model allows conditional heteroskedasticity and does not assume a functional form for the conditional expected returns. This is attractive because in models which assume a functional form, misspecification of the functional form can contaminate inferences about the number of common risk premiums. While our model is more general, it should also have improved power compared with other models. We illustrate how the approach can be further extended by examining models which allow limited (seasonal) fluctuations in conditional betas.

Using individual common stocks or using portfolios based on size rankings and industry affiliation, we reject models with fixed betas and a single risk premium. However, the latent variables models provide no evidence that more than two or three risk premiums are needed to capture expected returns over the 1927-87 period. Tests for the number of latent variables produce similar results in monthly and daily data. The similarity is interesting in view of a number of market microstructure considerations that should be more apparent in daily data. When we allow for conditional betas that can shift in January, the results for our general latent variable models are similar.

Our second contribution is a special case of our general latent

variable model. We refer to this as conditional mean variance spanning, as it extends the earlier work of Huberman and Kandel (1987). Huberman and Kandel examined the hypothesis that three size-based portfolios span the unconditional mean variance frontier of 30 size-based portfolios, and they found little evidence against this hypothesis. Our tests of conditional mean-variance spanning are more general because they do not assume homoskedasticity or normality. They examine the hypothesis that a particular subset of returns can generate the conditional mean-variance boundary for a given set of test assets.

Our tests produce evidence similar to Huberman and Kandel, when used to examine unconditional spanning for our sample of size-based common stock portfolios. However, when we condition on a common set of lagged information variables, we reject conditional mean variance spanning of the same sample. We also reject conditional beta pricing models, which correspond to mean variance intersection, using three or even four size portfolios as the factors.

Section I reviews latent variables models and presents a general model which does not assume that expected returns are linear. Section I also develops the tests of conditional mean variance spanning and discusses some econometric issues. Section II presents the empirical results. Section III summarizes and concludes.

I. Latent Variables Models

It is well known that mean-variance efficiency implies that expected returns can be described as a linear function of a single beta. The beta for each asset is measured relative to the mean-variance efficient portfolio

[e.g. Roll (1977)]. When expected returns and risks are conditioned on information about the state of the economy, then the betas, the expected risk premiums and the zero-beta rate may be time-varying. Latent variable models attempt to describe time-variation in expected returns using a small number of expected risk premiums, together with fixed beta coefficients.[2] The approach examines restrictions implied by an asset pricing model with the familiar form[3]

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

where λ_{ht} ≡ one of k unobserved ex post risk premiums or factor-mimicking portfolio returns;

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

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

At this level of generality, equation (1) does not assume that the risk factors are portfolio returns; they could be other state variables not observed by the econometrician [e.g. Cox, Ingersoll and Ross (1985)]. The model can generally be transformed, however, to one in which portfolios represent the factors [e.g. Breeden (1979)]. The asset betas, b_{ih}, will generally be time-varying functions of Z_{t-1}. The latent variables model of GF assumes that ratios of the b_{ih} are fixed parameters and specifies the dimension, K, of the asset pricing hypothesis. The expected risk premiums, $E(\lambda_{ht}|Z_{t-1})$, are assumed to be time-varying and unobserved, and are treated as the latent variables. We call the hypothesis that equation (1) holds with constant ratios of betas, for a given $K \geq 1$, the **"K Latent Variable**

Model."[4] If a K latent variable model is accepted, the results may be interpreted as indicating the number of "common factors," or time-varying risk premiums in the expected returns.[5]

Tests of latent variable models derive their power from the assumption that expected returns are changing over time and are correlated with observable instruments Z_{t-1}. GF assume that conditional expected returns given Z_{t-1} are linear with fixed coefficients, so that returns obey the regression model:

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

where Z_{t-1} is an L-vector of predetermined variables (including a constant), contained in the market's information set at time t - 1 and δ_i is the regression coefficient vector for each asset i, for i=0,...,N. We call equation (2) the **"Linear Expectations Assumption."** Given the linear expectations assumption, the following expression

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

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

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

The δ_j, for j = 0, ..., K, are the regression coefficients for K + 1 assets chosen as "**reference assets.**"[6] 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} are functions of the betas for assets i and j in equation (1), as described below. The restriction that the c_{ij} must sum to 1.0 for each asset i follows from the fact that $[1- \Sigma_{h=1}^{K} b_{ih}]$ is the coefficient on the "zero-beta" factor for asset i in equation (1). The information variables, 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 exceed 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 K latent variable hypothesis. The second is the linear expectations assumption. The third is the fixed residual covariance matrix assumption. Thus, the GF tests can be interpreted as examining:

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

Subsequent studies [e.g. Campbell (1987)] produce empirical evidence against the fixed covariance matrix assumption and extend the models to allow for conditional heteroskedasticity. We therefore concentrate our analysis on

the following two hypotheses:

$$H_1: \{ \text{ K latent variables, linear expectations } \}, \tag{6}$$

and

$$H_2: \{ \text{ K latent variables } \}. \tag{7}$$

To examine H_1, previous researchers relax the fixed covariance matrix assumption by estimating the restricted regression system (2), imposing the restrictions (4), using the generalized method of moments [GMM, see Hansen (1982)]. The parameters are chosen to minimize the quadratic form $Tg'Wg$, where $g=vec(u'Z/T)$ is the vector of sample orthogonality conditions, u is the T x n matrix of the error terms, Z is the T x L matrix of the instruments, T is the sample size, n is the number of equations and W is the inverse of the covariance matrix of the orthogonality conditions. This approach allows the covariance matrix of u to be conditionally heteroskedastic, and thus vary over time as a function of Z. The minimized value of the quadratic form is asymptotically distributed as a χ^2 variable under the null hypothesis, and provides a goodness-of-fit statistic. When the error terms are defined by the system (2), with the restrictions (4) imposed, the tests examine H_1 and we call the statistic the "**GMM1**" test statistic.[7]

A. Latent Variables Models with General Expected Returns

To examine H_2 we relax the assumption that expected returns are linear functions of $\mathbf{Z}_{t-1}$. This is accomplished by reformulating the model so that there exist constants c_{ij} such that:

$$R_{it} - \Sigma_{j=0}^{K} c_{ij} R_{jt} = \varepsilon_{it}, \tag{8a}$$

$$1 = \Sigma_{j=0}^{K} c_{ij}; \quad i = K+1, \ldots, N. \tag{8b}$$

$$E(\varepsilon_{it} \mid Z_{t-1}) = 0. \tag{8c}$$

The equations (8) may be derived as follows. Consider equation (1), where the $E(\lambda_{ht}|Z_{t-1})$; h=0,...,K are assumed to depend on Z_{t-1} but the functional form is not specified. In particular the expected premiums are not assumed to be linear functions of Z_{t-1}, nor are the individual asset returns assumed to be linear functions of Z_{t-1}.[8] Let $r \equiv \underline{R} - R_0$ be the vector of asset returns in excess of the return of the zero-th asset, R_0. The asset R_0 is not assumed to be a zero beta asset. Partition the vector of excess returns r as $r=(r_1', r_2')'$, where r_1 represents the first K reference assets and r_2 the remaining N-K test assets. Define the N x K matrix of excess betas, β, with typical element $\beta_{ij}=b_{ij}-b_{0j}$, and partition this matrix conformably with the excess returns as $\beta=(\beta_1', \beta_2')'$. Suppressing the time subscripts, equation (1) implies:

$$E(r_1|Z) = \beta_1 \; E(\lambda^*|Z) \tag{9a}$$

$$E(r_2|Z) = \beta_2 \; E(\lambda^*|Z), \tag{9b}$$

where $E(\lambda^*|Z)$ is the K-vector of expected risk premiums in excess of the expected zero beta return: $E(\lambda^*|Z) = (E(\lambda_{ht}|Z_{t-1}) - E(\lambda_{0t}|Z_{t-1}))$, h=1,...,K. If the reference assets are chosen so that the K x K matrix β_1 is invertible, we may use equation (9a) to solve for $E(\lambda^*|Z)$ in terms of $E(r_1|Z)$. Substituting this relation into (9b), and expressing the result in terms of

the original returns, we have:

$$E(\underline{R}_{2t}|Z_{t-1}) = [\underline{1}_{N-K} - \beta_2\beta_1^{-1}\underline{1}_K]\, E(R_{0t}|Z_{t-1}) + \beta_2\beta_1^{-1}\, E(\underline{R}_{1t}|Z_{t-1}) \; ; \qquad (10)$$

where $\underline{R}_{1t}$ in equation (10) is the K-vector of reference assets, $\underline{R}_{2t}$ are the N-K test assets, $\underline{1}_{N-K}$ is an N-K vector of ones and $\underline{1}_K$ is a K-vector of ones. If the error term, ε_{it}, is defined as in equation (8a), where the R_{it} are the N-K test assets and the R_{jt}, j=0,...,K are the zero-th asset and the other K reference assets, then equation (10) implies that there exist parameters c_{ij} such that the equations (8b) and (8c) are satisfied. In particular, the c_{ij} for i=K+1,...,N and for j=1,...,K are the elements of $\beta_2\beta_1^{-1}$ and the c_{i0}, for i=K+1,...,N are the elements of $[\underline{1}_{N-K} - \beta_2\beta_1^{-1}\underline{1}_K]$. The assumption that the matrix $\beta_2\beta_1^{-1}$ is fixed makes more precise our previous statements that the latent variable models assume that "ratios of the betas" are constant.

Note that 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. We use the general latent variable model of equation (8) to examine the hypothesis H_2.[9] Our tests are based on the GMM goodness-of-fit measure and we call the test statistic the **"GMM2"** test statistic.

B. Conditional Mean Variance Spanning and Intersection

Huberman and Kandel (1987) consider N-K regression equations for a set of asset returns, R_2, regressed on a smaller set of K+1 asset returns R_1. They show the restriction that the intercepts are zero and the regression

slopes sum to 1.0 for each asset is equivalent to the statement that the **unconditional** mean-variance boundary formed from the combined set of returns can be generated using only the subset of returns R_1. They refer to this as mean-variance spanning.

Beta pricing as in equation (1), where the assets R_1 are mimicking portfolios for the risk factors, is equivalent to mean-variance intersection. This is the hypothesis that only one combination of the assets R_1 is on the mean-variance boundary of the combined set of returns.

While spanning is a stronger hypothesis than intersection, Huberman and Kandel (1987) found little evidence against unconditional mean-variance spanning. It is therefore interesting to examine the effects of bringing conditioning information into the analysis.

Mean variance spanning is closely-related to our general latent variable model in equation (8). If we replace the orthogonality conditions in equation (8c) with the conditions: $E(\varepsilon_{it})=E(\varepsilon_{it}\ R_{jt})=0$, for $j=0,\ldots,K$, then the c_{ij} coefficients are the unconditional regression coefficients of the R_{it} on the R_{jt}. Tests of this system, suppressing an intercept, are tests of the hypothesis of unconditional mean-variance spanning. Using the GMM, these tests are more general than the tests examined by Huberman and Kandel (1987) and Lehmann and Modest (1988) in that they do not assume normality or homoskedasticity of the error terms.[10]

If we replace the orthogonality conditions of equation (8c) with the conditions that $E(\varepsilon_{it}|Z_{t-1})=E(\varepsilon_{it}\ R_{jt}|Z_{t-1})=0$, for $j=0,\ldots,K$, then the tests examine the hypothesis of **conditional** mean variance spanning. The tests assume that the conditional betas of the test assets on the spanning assets R_j, for $j=0,\ldots,K$, are fixed parameters. We implement our tests of

conditional mean-variance spanning using the GMM. The instruments are a vector of ones, the lagged variables in Z_{t-1}, and the spanning asset returns R_{jt}, for j=0,...,K.

Our general latent variable model, given by the equations (8), is modified to test for mean-variance spanning by simply changing the orthogonality conditions on the error terms. To test for mean-variance intersection, we introduce intercepts in equations (8a). The intercepts depend on the expected zero-beta rate, which introduces new parameter(s). The restrictions of (8b) are replace by the restriction that the intercepts are $(1-\Sigma_{j=0}^{K} c_{ij})$ multiplied by the expected zero beta return. Non-normal and heteroskedastic errors can be accommodated by using the GMM.[11]

Tests of mean variance spanning and intersection are related to the analysis of Hansen and Jagannathan (1991), who study the restrictions on an intertemporal marginal rate of substitution, m_t, implied by the general asset pricing expression $E\{m_t(1+R_{it})\}=1$, for all i. Mean variance spanning is equivalent to the restriction that a projection of the marginal rate of substitution on the vector of asset returns has coefficients equal to zero for all assets except the spanning assets R_{jt}, j=0,...,K, and for all possible values of an expected zero-beta rate. Intersection is equivalent to the same restriction holding for an unique value of the expected zero-beta rate. See Ferson (1992) for proofs and further discussion; see Cochrane (1992) for some interesting empirical tests.

C. Econometric Issues

Asset pricing models generally imply a condition that may be stated as $E(u|Z)=0$, where u is the model error term and the elements of Z are

predetermined public information. However, GMM estimation uses the weaker condition that $E(u'Z)=0$. Unless particular probability distributions for the asset returns are assumed (e.g., normality), the condition that u is uncorrelated with Z is weaker than the condition that $E(u|Z)=0$. Therefore, GMM tests of asset pricing models typically do not fully exhaust the implications of the asset pricing theories.[12] This is convenient for our tests of mean-variance spanning. Mean variance spanning does not imply that $E(\varepsilon_{it}|R_{jt})=0$ or that $E(\varepsilon_{it}|R_{jt},Z_{t-1})=0$, $j=0,\ldots,K$, so we do not wish to impose these conditions in our tests. These stronger conditions imply that the K+1 reference assets R_j are separating funds in the sense of Ross (1978) [see for example, Huberman and Kandel (1987), proposition 2].

The hypothesis H_2, as examined using the GMM2 test statistic, is less restrictive than H_1, as examined using the GMM1 test statistic. Logically, one expects to find less evidence in the data against H_2 than against H_1. For example, if the linear expectations assumption is misspecified then the error terms under H_1 can be related to the instruments and this can cause a rejection of the orthogonality condition $E(u'Z)=0$. Nonlinearity in the conditional mean returns is not a misspecification under H_2.

Tests of the general latent variable model using the GMM2 statistic should have higher power against some alternatives than should the tests of H_1 using the GMM1 statistic. This is because the power of the test statistic, for a given set of instruments Z and a given alternative for which $E(u'Z)$ is nonzero, is higher when the covariance matrix of the error term u is smaller. The power should be improved because system (8) can control more of the error variance than system (2). In system (2), the error variance is the portion of the asset return variance that is not explained

by the lagged instruments, which is large [see Table II]. In system (8), the contemporaneous returns R_j, j=0,...,K can "explain" a large fraction of the test asset returns, resulting in a smaller residual variance.[13]

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, as discussed above, and there could be finite sample problems which distort the results. The test statistics do not explicitly account for errors in the estimates of the covariance matrix of the orthogonality conditions, and these may be important in small samples. Finite sample problems should not be important in our large samples of daily data, but could be important in monthly sample sizes.

Ferson and Foerster (1992) provide simulation evidence on the finite sample properties of GMM tests of latent variable models. They find 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, especially in large models. Using an iterated approach, they find that the size of the tests are accurate for monthly samples with only half the numbers of observations in our samples. Following Ferson and Foerster (1992), we use an iterated GMM approach in our tests.[14]

II. Empirical Results

We begin our analysis with a brief look at the daily returns of the Dow Jones 30 (DJ30) common stocks, which were studied by Gibbons and Ferson (1985). This is an interesting sample because no studies have examined the more general latent variable models for individual common stocks. We then compare the results with an expanded sample of daily and monthly portfolio

returns.

Table I reports test results for the latent variable models of the DJ30. In the left-hand columns the instruments are a constant, the lagged return on a value weighted stock index and a dummy variable for Monday. These are the same as the instruments used by GF. In the right-hand columns, a dummy variable for the month of January (DJAN) is included as an additional instrument. In subperiods one through three, which approximates the time period used by GF, the GMM1 and GMM2 statistics produce no evidence against the single latent variable model. In the fourth subperiod (1980 to 1985) both test statistics reject (at the 0.05 level) a single latent variable. The tests provide robust evidence that more than a single time-varying risk premium is needed to capture the expected returns of the Dow Jones 30 stocks, as they do not assume that the residual covariance matrix is fixed (under H_1) or that the expected returns, conditional on the instruments, are linear with constant coefficients (under H_2).

Table I also reports tests for K=2 latent variable models and provides no evidence against the models under H_1 or H_2. The results are similar when the January dummy is excluded or included as an instrument. Neither the GMM1 nor the GMM2 statistics produce evidence against a two latent variable model for the daily returns of the DJ30 stocks.

A. Tests with Portfolio Returns

We examine daily and monthly returns for size and industry-grouped common stock portfolios. Results for each are summarized, but we focus on the monthly returns. Monthly returns are less susceptible to the biases from bid-ask effects and thin trading.[15] 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 can 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. "liquidity") demand shocks in order to focus on specialist behavior in the presence of asymmetric information, strategic trading and other such issues. These "exogenous" demands, however, should ultimately become endogenous. They may be driven by the same underlying factors that determine demands in standard equilibrium models of expected returns. We find it interesting, therefore, that our tests for the number of latent variables produce similar results for both daily and monthly returns.

The sample consists of common stocks of NYSE firms. Beginning in January of 1928 and ending in December of 1987, a total of 720 monthly observations are available. Ten 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 in each decile group. The daily size portfolio sample is similar, but the daily data begin 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.[16]

We also examine 12 portfolios of NYSE firms grouped by the same 2-digit SIC industry codes used in Breeden, Gibbons and Litzenberger (1989). 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 low grade bond portfolio returns are provided by Ibbotson Associates for 1928-1976 and by Blume, Keim and Patel (1991) for 1977 to 1987. The government bond returns are from CRSP.

B. Predicting Asset Returns

We include five predetermined instruments in our monthly tests. The motivation for including the 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 lagged one-month return of a three-month Treasury bill less the one-month return of a one-month bill. Campbell (1987) finds that a similar variable can predict monthly returns in both the bond and the stock markets.

D/P is the dividend yield of the CRSP value-weighted stock index, defined as the dividends paid on the index for the previous twelve months, divided by the current level of the index. A similar variable is studied by Rozeff (1984), Fama and French (1988b), Poterba and Summers (1988) and

others. Fama and French (1989) argue that dividend yields 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 studied by Fama and Schwert (1977), Breen, Glosten and Jagannathan (1989), Ferson (1989) and others.

DJAN equals one if the month is January and zero otherwise. We include the January dummy variable to capture seasonal patterns in returns [Rozeff and Kinney (1976), Keim (1983)] and for further analysis of possible seasonal changes in risk.

The predetermined variables used in the monthly regressions follow previous empirical work on predicting portfolio returns.[17] There is a natural concern, raised by a number of researchers, about predictability uncovered through collective "data snooping."[18] However, some evidence to support the view that the predictability is not spurious is available from studies using international data.[19] Furthermore, Ferson and Harvey (1991) find that stylized beta pricing models can "explain" much of the predictability uncovered by similar instruments in monthly post war U.S. data. They use a set of five economic risk factors, similar to Chen, Roll and Ross (1986). Some theoretical support for predictability using the lagged instruments is also available. For example, Bossaerts and Green (1989) develop a model in which conditional expected returns are inversely-related to the price of an asset. Kandel and Stambaugh (1990) develop model economies in which yields track time-varying expected risk premiums.[20]

Our monthly regression model takes the following form:

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

Table II summarizes the regression (11) for selected portfolios over the 1968-1987 subperiod. To conserve space, we report statistics only for the most recent subperiod. The coefficients of the regressions are similar to the findings of other studies over similar periods, so they are not shown. The adjusted R-squares of the regressions vary across the portfolios from less than 3 percent to over 19 percent. Recall that the latent variable models under H_1 [equation (4)] imply that the regression coefficients for all of the test assets are linear combinations of the coefficients for the reference assets. We conduct tests of various linear hypotheses on the coefficients, which suggest that the sample design should provide some power. The tests reject, for most of the instruments and subperiods, the hypotheses that the coefficients are jointly equal to zero or are equal across the portfolios. Similar results are found in the daily data.

In the right-hand column of Table II we report tests for heteroskedasticity in the residuals of regression (11). The test is that of White (1980), and is constructed by regressing the squares of the OLS residuals from (11) on the squares and cross products of the elements of the instrument vector, Z_{t-1}. The coefficient of determination from the regression, multiplied by the sample size, is asymptotically a chi-square variable and serves as the test statistic. The table reports the right-tail p-value from the chi-square distribution for each equation. We find strong evidence for heteroskedasticity in the equations for the bonds. We can also

reject homoskedasticity for the largest firms and for some of the industries, but the tests do not reject homoskedasticity for the smaller firms and many of the industries.[21]

The evidence of conditional heteroskedasticity suggests that tests of latent variables models under H_0 are problematic. Heteroskedasticity also suggests that the conditional betas of the test assets may be time-varying. However, latent variable models refer to betas with respect to unobserved risks factors. Risk factor-mimicking portfolios of the test assets will generally be formed with time-varying weights. Therefore, heteroskedasticity in the covariance matrix of the asset returns does not imply that a latent variable model with a small number of fixed betas is misspecified under H_1 or H_2.[22]

C. Tests of Latent Variable Models

Table III summarizes tests of the latent variable models using monthly data for the size and industry portfolios. The goodness-of-fit statistics are asymptotically chi-square variables, with degrees of freedom equal to the number of orthogonality conditions less the number of parameters. For the GMM1 (GMM2) statistic with N assets, K latent variables and L instruments, there are N×L ([N-K-1]×L) orthogonality conditions. There are [(K+1)×L + (N-K-1)×K] parameters in the model based on the equations (2) and (4), and there are K × (N-K-1) parameters in the model based on the equations (8).

Table III summarizes models with K=1,2 and 3 latent variables. The tests reject a single latent variable (K=1) under both H_1 and H_2 at standard significance levels for the full sample, except for the sample of industry

and bond portfolios under H_1. There is no evidence that more than two latent variables are required under H_1, using the GMM1 statistic. The GMM2 test statistic, examining H_2, produces smaller p-values in nearly every case than does the GMM1 test statistic, despite the weaker assumptions of H_2 compared with H_1. This is consistent with our view that the GMM2 test should have higher power than the GMM1 test.[23] The GMM2 test can reject the K=2 model for the overall sample at the 5% level. It can reject the K=3 model at the 10% level for the sample of size and bond portfolios, but not for the industry portfolio sample. However, the small p-values for the K=2 and K=3 models for the industry portfolios are driven by the second subperiod. The smallest p-value in the first and third subperiods for these cases is 0.068.

D. Conditional Mean-variance Spanning

Conditional mean-variance spanning is a stronger economic hypothesis than is a latent variable model. Therefore, if we reject a K-latent variable model we should also reject conditional mean-variance spanning using K+1 or fewer spanning portfolios. Given our evidence on the latent variable models, we conduct tests with K=2 (three spanning portfolios) and K=3 (four spanning portfolios).

We follow Huberman and Kandel (1987) by focusing on monthly data and using size-based portfolio returns as the spanning portfolios R_j, for $j=0,\ldots,K$ in system (8). Huberman and Kandel examine **unconditional** mean-variance spanning using three size-based portfolios (K=2). Their test assets are thirty size-based common stock portfolios.[24] They use an F test, assuming homoskedasticity and normality, and they find little evidence

against the spanning hypothesis for the 1964-1983 period.

Table IV summarizes the results of our tests for mean-variance spanning. The spanning portfolios for K=2 are the common stock size portfolios from deciles 1,5, and 10. When K=3, they are the size portfolios from deciles 1,4,7, and 10. Consistent with Huberman and Kandel, we find little evidence against unconditional spanning with K=2 (or K=3), when our remaining size-sorted portfolios are the test assets and the sample period is 1948-67 or 1968-87. However, we do reject unconditional spanning for 1928-47, a period not examined by Huberman and Kandel. (The p-values are 0.026 for K=2 and 0.036 for K=3.)

When we include bond returns in the test assets, or when the industry-grouped common stock portfolios are the test assets, the tests reject spanning (K=2 or K=3) in every subperiod. The evidence of Huberman and Kandel (1987) for size based portfolios as test assets, which seemed to support an unconditional spanning model using three size portfolios, does not extend to these alternative assets.[25] Since our tests do not rely on homoskedasticity or normality, the rejections are not attributed to problems with those assumptions.

The right-hand column of Table IV reports tests of conditional mean-variance spanning, using the same lagged variables that were used in the latent variables tests as the conditioning information. With conditioning information the evidence against spanning is stronger. The tests now produce evidence against conditional mean-variance spanning even when the seven remaining size portfolios are the test assets. The hypothesis is rejected in this case for the first subperiod (K=2 and K=3) and marginally in the third subperiod (K=2 only). The tests also reject spanning whenever the bonds or

industry portfolios are included as test assets.

We modify system (8) as described in section I.B. to test for conditional mean variance intersection. We assume for these tests that the expected zero beta return is a linear function of the instruments, and we estimate the coefficients of the linear function as additional parameters. The tests reject conditional intersection (p-values less than 1%) when the bonds are included as test assets with the size portfolios. In the second subperiod we reject conditional intersection for the industries. The p-values are 0.004 for the three portfolio model (K=2) and 0.037 for the four portfolio model (K=3). In the first and third subperiods the pvalues for the are 10% or larger when the industries are the test assets.

The evidence of Tables III and IV presents a challenge for future research. The latent variables models seem to hold out the hope that a small number of risk premiums may be able to explain the conditional expected returns. However, the spanning and intersection tests indicate that simply using size-based portfolios to proxy for the risk factors is not adequate.[26]

E. Extensions of the Tests

A number of empirical studies of conditional asset pricing models specify explicit functional forms for the conditional covariances or betas, which allow them to be time-varying functions of the instruments. Examples include Campbell (1987) and Ferson (1989) in latent variable models and Harvey (1989) and Shanken (1990), who specify the risk factors.[27] The models in this paper can also be extended to allow for time-varying conditional betas and beta ratios (the c_{ij} coefficients). The generalization is accomplished by replacing the fixed c_{ij}'s in the models

with the specified function, which may depend on the instruments Z_{t-1} or other variables. We illustrate this idea by allowing the c_{ij} coefficients to shift in January. Such a model is motivated by earlier work of Ferson and Keim (1984), Rogalski and Tinic (1986), Keim and Stambaugh (1986) and Shanken (1990) who find evidence that betas may shift in January.

We first conduct tests of the hypothesis that expected risk premiums are zero outside of January, as was suggested by Tinic and West (1984). If this is the case we expect that tests which allow betas to shift in January should have low power. We form the differences between the returns of each portfolio and the first asset in each system of equations. We modify the predictive regressions as follows. The return differences are regressed on a constant, the predetermined instruments (excluding the January dummy), and the instruments multiplied by the dummy variable for January. We test the hypothesis that the coefficients on the variables without the dummy are jointly equal to zero. A heteroskedasticity-consistent Wald test strongly rejects the hypothesis that the non-January premiums are zero for each sample of assets in every subperiod, both in the daily and the monthly data. The 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. Because these tests are not dependent on specific proxies for the underlying risk factors, they complement the recent evidence of Shanken (1990).[28]

Using the modified predictive regressions, we generalize our tests of latent variable models and conditional mean variance spanning to allow the c_{ij}'s to shift in January.[29] Assume that the beta coefficients in January

(b^{J}_{ij}), are possibly different from betas during the rest of the year (b_{ij}). The c_{ij} are replaced by the expression: $\{c^{NJ}_{ij} + (c^{J}_{ij} - c^{NJ}_{ij})DJAN_t\}$, where $DJAN_t$ equals 1.0 if the return being forecast occurs in a January and 0.0 otherwise. We therefore estimate separate relative beta coefficients for January (c^{J}_{ij}) and for the other months (c^{NJ}_{ij}). The tests otherwise proceed as before. The results for monthly data are summarized in Table V. Latent variable models are shown in panel A and the tests of conditional mean variance spanning are in panel B.

Recall the GMM1 tests of H_1 that were reported in Table III could not reject the latent variable models (with one exception). Allowing for January effects in the betas, we find that the GMM1 statistic is unable to reject any of the models. With the larger number of parameters in the regression model for the expected returns, the GMM1 statistic indicates that the latent variable models' restrictions impose virtually no additional structure on the regressions. (The p-values would all appear as 0.999, so we do not report them in the table.)

Panel A of Table V reports tests of H_2 using the GMM2 statistic. Recall that under H_2 we do not assume a functional form for conditional expected returns and we allow for conditional heteroskedasticity. Although the numbers differ from those in Table III, where we assumed constant betas, the overall inferences implied by the GMM2 tests are similar.

Panel B of Table V examines conditional mean variance spanning using the model of seasonal changes in beta. The results may be compared with those of Table IV, where the betas were held fixed. The overall results are similar. We reject spanning (K=2 or K=3) of the size portfolio sample only in the first subperiod, which is 1928-1947. When we introduce the bonds as

additional test assets, or when we study the sample of industry and bond portfolios, we reject spanning in the overall samples. However, there are cases in which we do not reject spanning for these samples in the subperiods (one for K=2 and three cases when K=3), which shows that the spanning model is sensitive to allowing January shifts in the betas.

III. Concluding Remarks

This paper studies the behavior of conditional expected returns over time on common stocks and bonds. We extend the latent variables method of Gibbons and Ferson (1985) to allow conditional heteroskedasticity and relax a common assumption that conditional expected returns are linear functions of predetermined instruments. A special case of this model, which we denote as conditional mean variance spanning, generalizes earlier work of Huberman and Kandel (1987). We show how to extend the models to relax the assumption that conditional betas are fixed parameters by replacing that assumption with a functional form for the betas. This is illustrated by examining models in which the beta exhibits a January seasonal.

We examine an extensive sample of daily returns, monthly returns and instrumental variables for the predetermined conditioning information. Gibbons and Ferson (1985) did not reject a single latent variable model for the Dow Jones 30 common stocks, but we find that such a model can be rejected using an extended sample for the Dow Jones 30, and for samples of size and industry portfolios. These results are robust to heteroskedasticity and to any assumption about the functional form of conditional expected returns. However, the tests provide little evidence for more than two or three latent variables in the time-varying expected

returns.

On the one hand, our results lead us to be optimistic about the potential ability of linear asset pricing theories with a small number of common factors to capture the predictable variation of security returns both across assets and over time. A small number of common factors in the expected returns seems to be indicated, and interestingly, the results for monthly and daily returns are similar. However, our tests which specify size-based common stock portfolios as proxies for the risk factors can reject models using three or even four portfolios. Therefore, the search for a small number of variables which can capture conditional expected asset returns should continue.

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, 1991, Returns and volatility of low-grade bonds, 1977-1989, Journal of Finance 46, 49-74.

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

Bodurtha, James N. and Nelson C. Mark, 1991, Testing the CAPM with time-varying risks and returns, Journal of Finance 46, 1485-1505.

Bollerslev, Tim ,1986, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics 31, 307-327.

Bollerslev, Tim, Robert Engle and J. Wooldridge, 1988, A capital asset pricing model with time-varying covariances, Journal of Political Economy 96, 113-131.

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.

Braun, Phillip, Dan Nelson and Alan Sunier, 1991, Good news, bad news, volatility and betas, working paper, University of Chicago.

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.

Breen, William, Lawrence Glosten and Ravi Jagannathan, 1989, Economic significance of predictable variations in stock index returns, Journal of Finance 44, 1177-1190.

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

Campbell, J. and Yasushi Hamao, 1992, Predictable bond and stock returns in the United States and Japan: A study of long-term capital market integration, Journal of Finance 47, 43-69.

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

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

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

Cox, John C., Jonathan E. Ingersoll, Jr. and Stephen A. Ross, 1985, A theory of the term structure of interest rates, Econometrica 53, 385-407.

Cutler, D. M., J. M. Poterba and L. H. Summers, 1988, International evidence on the predictability of stock returns," working paper, Massachusetts Institute of Technology.

Engle, Robert F., 1982, Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation, Econometrica 50, 987-1008.

Evans, Martin D.D., 1991, Expected returns, time-varying risk and risk premia, working paper, New York University.

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, Eugene F. and James D. MacBeth, 1973, Risk, return and equilibrium: empirical tests, Journal of Political Economy 81, 607-636.

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

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

Ferson, W. E., 1992, Theory and empirical testing of asset pricing models, Chapter A6, forthcoming in Robert A. Jarrow, William T. Ziemba and Vojislav Maksimovich, eds: The Finance Handbook (North Holland publishers).

Ferson, W. E. and Foerster, S. R., 1992, Finite sample properties of the Generalized Method of Moments in tests of conditional asset pricing models, working paper, University of Chicago and University of Western Ontario.

Ferson, W. E. and C. Harvey, 1991, The variation of economic risk premiums, Journal of Political Economy 99, 385-415.

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, Phd dissertation, The Wharton School of the University of Pennsylvania.

Foerster, S. R. and D. Porter, 1991, Calendar- and size-based anomalies in canadian stock returns, working paper, University of Western Ontario.

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

Hansen, Lars P., 1982, Large sample properties of generalized method of moments estimators, Econometrica 50, 1029-1054.

Hansen, Lars P. and Ravi K. Jagannathan, 1991, Implications of security market data for models of dynamic economies, Journal of Political Economy 99, 225-262.

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

Harvey, Campbell, R., 1989, Time-varying conditional covariances in tests of asset pricing models, Journal of Financial Economics 24, 289-317.

Harvey, Campbell R., 1991, The world price of covariance risk, Journal of

Finance 46, 111-158.

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

Jagannathan, Ravi K., 1983, Three essays on the pricing of derivative claims, Phd dissertation, Carnegie Mellon University.

Kandel, Shmuel A. and Robert F. Stambaugh, 1990, Expectations and volatility of consumption and asset return, Review of Financial Studies 3, 207-232.

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. and David M. Modest, 1988, The empirical foundations of the arbitrage pricing theory, Journal of Financial Economics 21, 213-254.

Lim, K.G., 1985, Estimating and testing rational expectations models under non-stationarity, Phd dissertation, Stanford University.

Lo, Andrew W. and A. Craig MacKinlay, 1990a, Data snooping and biases in tests of financial models, Review of Financial Studies 3, 431-468.

Lo, Andrew W. and A. Craig MacKinlay, 1990b, An econometric analysis of nonsynchronous trading, Journal of Econometrics 45, 181-211.

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

MacKinlay, A. Craig and Matthew Richardson, 1991, Using the generalized method of moments to test mean-variance efficiency, Journal of Finance 46, 511-528.

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.

Nelson, Dan 1991, Conditional heteroskedasticity in asset returns: a new approach, Econometrica 59, 347-370.

Ng, Lilian, 1991, Test of the CAPM with Time-varying covariances: a multivariate GARCH approach, Journal of Finance 46, 1507-1521.

Porter, D., 1992, The probability of a trade at the ask: an examination of interday and intraday behavior, Journal of Financial and Quantitative Analysis (forthcoming).

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. A., 1978, Mutual fund separation in financial theory: the separating distributions, Journal of Economic Theory 17, 254-83.

Rozeff, Michael, 1984, Dividend yields and equity risk premiums, Journal of Portfolio Management, Fall, 68-75.

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

Shanken, Jay, 1986, Nonsynchronous data and the covariance/factor structure of returns, Journal of Finance

Shanken, Jay, 1990, Intertemporal asset pricing: an empirical investigation, Journal of Econometrics 45, 99-120.

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

Solnik, Bruno, 1991, Using the predictability of foreign asset returns, working paper, HEC-School of Management.

Stambaugh, Robert F., 1983, Arbitrage pricing with information, Journal of Financial Economics 12, 357-370.

Stambaugh, Robert F., 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 versus 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.

FOOTNOTES

1. See Ferson (1992) for a recent review.

2. An alternative approach models changing conditional betas directly. Fama and MacBeth (1973), Campbell (1987), Shanken (1990), Bodurtha and Mark (1991), Ng (1991), Ferson and Harvey (1991), Braun, Nelson and Sunier (1991) and the last section of this paper provide examples.

3. Examples of asset pricing models like equation (1) include those of Sharpe (1964), Black (1972), Merton (1973), Long (1974), Breeden (1979) and Cox, Ingersoll and Ross (1985).

4. See Stambaugh (1983) for a conditional version of the Arbitrage Pricing Model (APT) which implies that a K latent variable model holds approximately in large markets. See Connor and Korajczyck (1989) for an equilibrium model in which conditional betas are constant over time. The assumption that ratios of the betas are constant will be made more precise below.

5. An alternative interpretation is to view the tests as indicating the behavior of conditional covariances of returns with a benchmark pricing variable. See Gibbons and Ferson (1985), Campbell (1987), Ferson (1989) and Wheatley (1989) for discussions.

6. The reference assets must be chosen so that the matrix of their betas and a unit vector is nonsingular; that is, they must span the K risk factors and they 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 and Foerster (1992)].

7. Hansen (1982) shows that sufficient conditions to apply the GMM include the assumption that the data are strictly stationary and ergodic. Jagannathan (1983) and Lim (1985) show how to accommodate seasonality. We assume that the data satisfy the conditions needed to apply the GMM in all of our tests.

8. In contrast, if the linear expectations assumption holds for the asset returns and if the b_{ih} are fixed parameters, then the K latent variable hypothesis implies that the expected risk premiums must also be linear functions.

9. Ferson (1990) and Harvey (1991) examine a special case of system (8) stated for excess returns, in which equation (8b) is not imposed.

10. The R_{jt}, j=0,...,K, plus a constant term, are the instrumental variables. The system is overidentified because the intercepts are suppressed and the c_{ij}'s are constrained to sum to 1.0. See MacKinlay and Richardson (1991) for related tests of unconditional mean variance efficiency using the GMM.

11. Shanken (1990) provides regression tests of beta pricing which do not assume normality and which also allow heteroskedasticity. Shanken assumes that conditional betas are linear functions of the instruments.

12. For a given set of instruments Z, $E(u|Z)=0$ implies that $E(u \times f(Z))=0$ for all functions f(.). Typically, studies use only one f(.) function.

13. This is the case in our sample even though, in the general latent variable model, equation (8a) is not a regression in which the error terms must be uncorrelated with the R_j's. We compute the determinant of the covariance matrix of the error terms in equation (8), divided by the determinant for the error terms of (2). If the ratio is less than one the generalized error variance is smaller in the GMM2 than in the GMM1. For the six subperiods and asset samples described in Table III we find that the ratio varies from 0.11 to 0.96, and the average value is 0.63.

14. Specifically, we construct the weighting matrix W using the parameter estimates from the m-th stage minimization, use this matrix to find parameters for stage m+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.

15. 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 (1992) 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 may influence regression models that contain lagged market returns [Reinganum (1982), Lo and MacKinlay (1990b)]. The empirical evidence in Shanken (1987) raises questions about the use of daily data when estimates of security covariances are important. We conducted some analysis (available by request) which suggested that these are not serious problems in our daily sample for the DJ30. However, the biases are likely to be important for low-price, thinly-traded stocks that are concentrated in portfolios of small firms.

16. 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 (1991) study the effectiveness of such an approach in reducing the bias in measured portfolio returns. Their evidence suggests that the approximation is accurate enough for our purposes.

17. 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 period. Our results for more recent data therefore provide out-of-sample evidence.

18. Such concerns are raised by Merton (1985) and Lakonishok and Smidt (1988) and are analyzed by Lo and MacKinlay (1990b).

19. Cutler, Poterba and Summers (1988), Solnik (1991) and others find that dividend yields have predictive power for future stock returns in other countries. Campbell and Hamao (1992) find that predictable components of bond and stock returns are highly correlated between the U.S. and Japan. Harvey (1991) finds that a set of lagged instruments similar to ours has predictive power for stock returns in other countries.

20. We have replicated many of our tests using alternative choices for the monthly instruments. We replaced the dividend yield variable with a detrended price level variable. The price level is detrended by dividing the end of month level of the index into the average level over the previous year, similar to Keim and Stambaugh (1986). 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.

21. These conclusions are similar to the evidence reported in Shanken (1990), who uses a different set of instruments and monthly data for 1953-1982. We find stronger evidence against homoskedasticity in our first subperiod, where the largest p-value is 0.085 (leisure industry). Results for the second subperiod are similar to those reported for the third subperiod in Table II. Of course, White's (1980) test only detects heteroskedasticity that is linearly related to the levels and cross products of the instruments, and is therefore of limited power.

22. For direct evidence on the behavior of conditional betas when the risk factors are specified see Keim and Stambaugh (1986), Ferson (1990), Shanken (1990) and Braun, Nelson and Sunier (1991). Ferson and Harvey (1991) and Evans (1991) allow time-varying betas and conclude that the variation in betas contributes little to the variation in the expected returns of size and industry-grouped portfolios in post-war monthly data.

23. Recall that a smaller number of equations are also involved for a given asset sample and number of latent variables in the GMM2 test. This should also affect the power of the tests.

24. Their sample design differs from ours in several ways. They first form 33 size-based portfolios. An equally-weighted combination of the first eleven forms the first spanning portfolio, the next eleven comprise a second and the last eleven form the third. To avoid singularities, they then use only 30 of the original 33 size portfolios as their test assets.

25. Lehmann and Modest (1988) conduct tests of unconditional spanning using APT factors, which are portfolios of stocks constructed by factor analysis. They examine test portfolios grouped on size, dividend yield and own variance using data for 1963-82. They also find that their spanning models are rejected.

26. This is consistent with the results of Ferson (1990), who examined quarterly data and used a different approach to specify the risks factors. Shanken (1990) rejects a model where a government bond and a value-weighted stock index are the factors. However, Ferson and Harvey (1991) report that a model with five economic factors can explain much of the predictable time-variation in post war monthly returns of size and industry portfolios.

27. Harvey (1989) assumes that conditional covariances with a market index are ratios of linear functions. Autoregressive conditional heteroskedasticity models [ARCH, see Engle (1982)] and their generalizations [e.g., GARCH, see Bollerslev (1986) and EGARCH, see Nelson (1991)] have also been used to model conditional second moments when the risk factors are specified [e.g. Bollerslev, Engle and Wooldridge (1988), Bodurtha and Mark (1991), Ng (1991)].

28. Shanken specifies the risk factors and documents nonzero expected risk premiums outside of January, using monthly data for 1953-1982.

29. While there are other formal justifications, one motivation for this regression model is to assume that the underlying unobserved expected risk premiums are linear with fixed coefficients (which do not shift in January) and that the assets betas shift in January as indicated above.

TABLE I

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:

$$R_1 = Z\delta_1 + \varepsilon_1$$

$$R_2 = Z\delta_1 C + \varepsilon_2,$$

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

where $R=(R_1,R_2)$ is a T x (N+1) matrix of daily returns, partitioned by columns, with K+1 columns in R_1 and N-K columns in R_2. Z is a T x L matrix of predetermined instrumental variables, and ℓ is a vector of ones. δ_1 and C are L x (K+1) and (K+1) x (N-K) matrices of parameters. 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 time t values of the error terms in $u=(\varepsilon_1,\varepsilon_2)$ have conditional mean zero given the time t-1 instruments in Z. The orthogonality condition tested is $E(u'Z)=0$. The GMM2 statistic is the value of the criterion function for the reformulated model:

$$u = R_2 - R_1 C,$$

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

In the left-hand columns the instruments Z are a constant, the lagged return on a value weighted stock index and a dummy variable for Monday. In the right-hand columns 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.

		A. Instruments exclude January Dummy:		B. Instruments Include January Dummy	
No. Latent Vars.	Subperiod	GMM1	GMM2	GMM1	GMM2
1	1963-68(T=1478)	0.362	0.291	0.497	0.436
1	1969-73(T=1262)	0.675	0.268	0.703	0.344
1	1974-79(T=1518)	0.565	0.350	0.329	0.077
1	1980-85(T=1519)	0.036	0.030	0.096	0.005
1	1963-85[a]	0.221	0.048	0.281	0.011
2	1963-68	0.987	0.987	0.976	0.614
2	1969-73	0.940	0.907	0.890	0.839
2	1974-79	0.802	0.724	0.681	0.257
2	1980-85	0.980	0.972	0.991	0.862
2	1963-85[a]	0.999	0.999	0.999	0.814

[a] The aggregate test statistic sums the chi-square values across the subperiods.

TABLE II

Predictive regression results for monthly rates of return for size- and industry-grouped common stock portfolios. The data are for 1968-1987 (240 observations). 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. D/P is the dividend yield of the CRSP value-weighted stock index. TB is the nominal, one-month Treasury bill rate. DJAN is a dummy variable equal to one if month t is January and zero otherwise. adj. R^2 is the adjusted R-square, ρ_1 is the first order autocorrelation of the regression residual, and χ^2(pv) is the right-tail probability value for White's (1980) chi-square test for conditional heteroskedasticity of the regression residual.

Results for the regression model:

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

Portfolio[a]	adj. R^2	ρ_1	χ^2(pv)
Decile 1	19.0%	-0.04	.469
Decile 2	18.5	0.01	.376
Decile 5	14.3	0.03	.306
Decile 10	7.4	-0.04	.011
Petroleum	3.7	-0.01	.000
Fin/RE	6.3	0.08	.004
Cons. Dur.	12.3	0.01	.363
Basic Ind.	5.8	-0.04	.099
Food/Tob.	6.0	0.06	.001
Constr.	10.9	0.02	.179
Cap. Goods	10.4	0.00	.203
Trans.	10.1	0.01	.078
Utilities	6.4	0.08	.002
Textiles/Trade	6.4	0.13	.479
Services	13.1	0.06	.047
Leisure	11.5	0.09	.276
Govt. Bond	2.7	0.02	.000
Junk Bond	9.1	0.14	.000

[a] 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 III

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 T x (N+1) matrix of monthly returns, partitioned by columns, with K+1 columns in R_1 and N-K columns in R_2. Z is a T x L matrix of predetermined instrumental variables, and ℓ is a vector of ones. δ_1 and C are L x (K+1) and (K+1) x (N-K) matrices of parameters. 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 time t values of the error terms in $u=(\varepsilon_1,\varepsilon_2)$ have conditional mean zero given the time t-1 instruments in Z. The orthogonality condition tested is $E(u'Z)=0$. The GMM2 statistic is the value of GMM criterion function for the reformulated model:

$$u = R_2 - R_1 C,$$

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

The instruments are a constant, the dividend yield of the CRSP value-weighted stock index, the level of the one-month treasury bill, the lagged excess return 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 assets are ten value-weighted, size-based common stock portfolios, twelve industry-grouped portfolios, a long-term government bond (GB) and a low-grade corporate bond (junkret). Each subperiod has 240 monthly observations. There are 720 observations in the 1928-1987 sample period. The right-tail probability values for the test statistics are reported in the table.

table III page 2

No. Latent Variables	Subperiod	Assets	GMM1	GMM2
1	1928-47	10 Size, GB, Junkret	0.195	0.089
1	1948-67		0.039	0.024
1	1968-87		0.109	0.101
1	1928-87[a]		0.015	0.005
1	1928-47	12 Industry, GB, Junkret	0.630	0.638
1	1948-67		0.305	0.007
1	1968-87		0.257	0.134
1	1928-87[a]		0.333	0.027
2	1928-47	10 Size, GB, Junkret	0.728	0.510
2	1948-67		0.097	0.049
2	1968-87		0.434	0.059
2	1928-87[a]		0.313	0.032
2	1928-47	12 Industry, GB, Junkret	0.726	0.610
2	1948-67		0.113	0.001
2	1968-87		0.885	0.342
2	1928-87[a]		0.811	0.025
3	1928-47	10 Size, GB, Junkret	0.768	0.447
3	1948-67		0.242	0.105
3	1968-87		0.687	0.068
3	1928-87[a]		0.635	0.054
3	1928-47	12 Industry, GB, Junkret	0.863	0.745
3	1948-67		0.171	0.003
3	1968-87		0.939	0.639
3	1928-87[a]		0.813	0.105

[a] The aggregate test statistic sums the chi-square values across the subperiods.

TABLE IV

Tests of Conditional and Unconditional Mean-variance Spanning.

The model is:

$$\eta = R_2 - R_1 C,$$

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

where $R=(R_1,R_2)$ is a T x N matrix of monthly returns, ℓ is a vector of ones and C is a $(K+1)x(N-K-1)$ matrix of parameters. K is the dimension of the asset pricing model (i,e., the number of spanning assets minus one) and N is the total number of assets; there are N-K-1 test assets. The test 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 η is uncorrelated with the (K+1) spanning returns R_1. When K=2 the spanning returns are the size portfolios from the smallest, largest and 5-th decile, a total of 3 returns. When K=3 the portfolios of size deciles 1,4,7, and 10 are used. When unconditional spanning is tested, the orthogonality condition is $E(\eta)=E(\eta' R_1)=0$. When conditional mean variance spanning is tested, the orthogonality condition is $E(\eta'[R_1,Z])=0$, where Z is a T x L matrix of predetermined instrumental variables. The instruments are a constant, the dividend yield of the CRSP value-weighted stock index, the level of the one-month treasury bill, the lagged excess return 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 assets are ten value-weighted, size-based common stock portfolios, twelve industry-grouped portfolios, a long-term government bond (GB) and a low-grade corporate bond (junkret). Each subperiod has 240 monthly observations. There are 720 observations in the 1928-1987 sample period. The right-tail probability values for the test statistics are reported in the table.

table IV page 2

Panel A: K=2 (Three spanning portfolios)

Subperiod	Test Assets	Unconditional Spanning (p-values)	Conditional Spanning (p-values)
1928-47	7 Size Portfolios	0.026	0.010
1948-67		0.159	0.141
1968-87		0.516	0.065
1928-87[a]		0.048	0.003
1928-47	7 Size, GB, Junkret	0.000	0.003
1948-67		0.000	0.000
1968-87		0.000	0.001
1928-87[a]		0.000	0.000
1928-47	12 Industry Portfolios	0.001	0.022
1948-67		0.000	0.000
1968-87		0.000	0.001
1928-87[a]		0.000	0.000

Panel B: K=3 (Four spanning portfolios)

Subperiod	Test Assets	Unconditional Spanning (p-values)	Conditional Spanning (p-values)
1928-47	6 Size Portfolios	0.036	0.022
1948-67		0.117	0.295
1968-87		0.930	0.515
1928-87[a]		0.127	0.072
1928-47	6 Size, GB, Junkret	0.000	0.009
1948-67		0.000	0.000
1968-87		0.001	0.001
1928-87[a]		0.000	0.000
1928-47	12 Industry Portfolios	0.001	0.041
1948-67		0.000	0.002
1968-87		0.000	0.005
1928-87[a]		0.000	0.000

[a] The aggregate test statistic sums the chi-square values across the subperiods.

TABLE V

The effects of allowing conditional betas to shift in January.

The first panel summarizes models with K=1 and 2 latent variables. The latent variable model is:

$$R_1 = Z^* \delta_1 + \varepsilon_1$$

$$R_2 = Z^* \delta_1 C_t + \varepsilon_2,$$

$$\ell' C_t = \ell,$$

$$C_t = C^{NJ} + (C^J - C^{NJ})\ DJAN_t,$$

where $R=(R_1,R_2)$ is a T x N matrix of monthly returns, Z^* is the T x 2(L-1) matrix (Z,Z x DJAN), where DJAN is a dummy variable for the month of January and Z is the T x (L-1) matrix of instruments consisting of a constant, the dividend yield of the CRSP value-weighted stock index, the level of the one-month treasury bill, the lagged excess return of a three-month over a one-month bill, and the lagged return of the CRSP equally-weighted stock index. ℓ is a vector of ones. δ_1, C^{NJ} and C^J are parameters. The GMM2 statistic is the value of the GMM criterion function for the reformulated model:

$$\eta = R_2 - R_1\ C_t,$$

$$\ell' C_t = \ell,$$

$$C_t = C^{NJ} + (C^J - C^{NJ})\ DJAN_t.$$

The orthogonality condition is $E(\eta'(Z,DJAN))=0$. The tests of mean variance spanning in panel B use the same system of equations, where R_1 is the T x (K+1) matrix of spanning assets, K is the dimension of the asset pricing model and N is the total number of assets. There are N-K-1 test assets in the vector R_2. When K=2 the spanning assets are the size portfolios from the smallest, largest and 5-th decile, a total of 3 returns. When K=3 the portfolios of size deciles 1,4,7, and 10 are used. The test statistic is the minimized value of the GMM criterion function for the system, based on the orthogonality condition $E(\eta'[R_1,Z,DJAN])=0$. The assets are ten value-weighted, size-based common stock portfolios, twelve industry-grouped portfolios, a long-term government bond (GB) and a low-grade corporate bond (junkret). Each subperiod has 240 monthly observations. There are 720 observations in the 1928-1987 sample period. The right-tail probability values for the test statistics are reported in the table.

table V page 2

Panel A: Latent Variable Models			(GMM2)
No. Latent Varis.	Subperiod	Assets	
1	1928-47	10 Size, GB,	0.314
1	1948-67	Junkret	0.025
1	1968-87		0.034
1	1928-87[a]		0.008
1	1928-47	12 Industry,	0.582
1	1948-67	GB, Junkret	0.006
1	1968-87		0.027
1	1928-87[a]		0.006
2	1928-47	10 Size, GB,	0.590
2	1948-67	Junkret	0.142
2	1968-87		0.015
2	1928-87[a]		0.037
2	1928-47	12 Industry,	0.179
2	1948-67	GB, Junkret	0.003
2	1968-87		0.721
2	1928-87[a]		0.023

Panel B: Conditional Mean-variance Spanning			
		K=2	K=3
Subperiod	Test Assets	(3 spanning portfolios)	(4 spanning portfolios)
1928-47	7 Size Portfolios	0.035	0.042
1948-67		0.158	0.380
1968-87		0.250	0.812
1928-87[a]		0.025	0.234
1928-47	7 Size, GB,	0.004	0.002
1948-67	Junkret	0.000	0.005
1968-87		0.008	0.333
1928-87[a]		0.000	0.000
1928-47	12 Industry	0.172	0.609
1948-67	Portfolios	0.001	0.010
1968-87		0.006	0.119
1928-87[a]		0.000	0.028

[a] The aggregate test statistic sums the chi-square values across the three subperiods.

[b] The overall p-value is for the first two subperiods only.

[nc] Indicates that we were unable to obtain satisfactory convergence of the algorithm.