Multivariate tests that ask if the α's are jointly different from each other

Wentworth Boynton*

April 4, 2009

Abstract

The standard multivariate test ranks stocks by a characteristic, such as market size or past performance, and forms a set of dependent-variable portfolios. Portfolio returns are regressed on a risk-based model, and the test asks if the intercepts are jointly different from zero. The paper provides an alternative test that asks if the intercepts are jointly different from each other. The test subtracts the returns of closely-ranked portfolios. Examples show that smaller generalized variance and better power than standard tests.

Key words: asset-pricing tests, performance measurement

JEL Classification: G11, G12

---

*Department of Economics and Finance, University of New Haven, 300 Boston Post Road, West Haven, CT 06516; email: cboynton@newhaven.edu.

1 Introduction

The central idea of asset pricing is that the covariance of the returns of risky portfolios are priced relative to the marginal utility of wealth (Breedon, 1979, Lucas, 1978, and Merton, 1973). In these tests, one typically sorts the entire stock market by a characteristic, such as market-size or past performance, forms portfolios, and regresses these portfolios on some chosen risk-based model. Let $\hat{\boldsymbol{\alpha}}$ be the estimated performance vector and $\hat{\boldsymbol{\Sigma}}$ the estimated error disturbance matrix. The quadratic form $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\alpha}}$ summarizes the statistical distance between the tangency portfolio and the factor pricing model. Gibbons, Ross, and Shanken (1989) (GRS hereafter) study the small-sample properties of these tests.[1]

The chief goal of the paper is to provide a specialized measue for the test of whether the performance measures are different from each other. This test is a stricter than the standard multivariate test that asks if the performance measures are jointly different than zero. It is well known that standard tests might reject for a variety of reasons. For example, tests might show mispricing across all the dependent-variable portfolios but little differences in pricing within these portfolios. The proposed test is useful for evaluating if there is a useful trade rule to sell short one portfolio and buy another.

Let $\mathbf{1}_N$ be a $N \times 1$ vector of ones and $\mathbf{I}_N$ be a $N \times N$ identity matrix. Assume $\hat{\boldsymbol{\Sigma}} = \mathbf{I}_N \sigma^2$. Following GRS, define $\mu_1 = \frac{1}{N}\mathbf{1}'_N\hat{\boldsymbol{\alpha}}$, the average α, and $\mu_2 = \frac{1}{N}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\alpha}}$, the average $\hat{\alpha}^2$. GRS write, "One can view μ_1 as a measure of the "average" misspecification across assets while μ_2 indicates the noncentral dispersion of the departures from the null ...." If $\mu_1 \rightarrow 0$, the standard test has no common mispecification ($\mu_1 = 0$), μ_2 tests if the α's are different from each other. Unfortunately, there are two barriers to implementing a test where the α's are jointly different from each other. First, one must find a pricing factor that generates no common mispecification; in light of the Roll critique, such a factor setup is problematic. Second, the setup assumes that the $\hat{\boldsymbol{\Sigma}}$ is diagonal and constant across its diagonal terms; these diagonal conditions rarely hold.

An alternative approach is to subtract returns from closely ranked portfolios. More formally, let $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_2, ..., \mathbf{y}_N$ be the set of dependent-variable portfolios. Compute the difference vectors $\mathbf{d}_1 = \mathbf{y}_1 - \mathbf{y}_2, \mathbf{d}_2 = \mathbf{y}_2 - \mathbf{y}_3, \mathbf{d}_3 = \mathbf{y}_3 - \mathbf{y}_3, ..., \mathbf{d}_{N-1} = \mathbf{y}_{N-1} - \mathbf{y}_N$. The multivariate

[1]The GRS paper is preceded by a number of papers that explored some of these same issues. See for example Gibbons (1982), Jobson and Korkie (1982), MacKinlay (1987), and Stambaugh (1982).

tests on this set of dependent variables asks whether the α's are different from each other, conditional on a risk-based model. Each of the $\mathbf{d}_i$ portfolios are "absolute return" strategies that one hypothesizes will yield abnormal profits from the trade rule. If the closely-ranked portfolio returns closely covary, it is likely that the test will have a substantially smaller generalized variance and better power. The paper calls this design the "difference-return test." For empirical tests, the paper uses momentum and market-size sorts to organize the dependent-variable portfolios and use CAPM and Fama and French 3-factor risk-based models.

Several papers use principal components for constructing a factor structure for pricing risk (Connor and Korajczyk, 1986, 1988, 1995; and Grinblatt and Titman, 1985). Principal component tests seek sets of factors that purge the system of common stock return comovements. Tests assume that the error covariance matrix $\hat{\mathbf{\Sigma}}$ is approximately diagonal. In contrast, most research assumes a specific risk-based model a priori, such as CAPM or the Fama and French (1993) 3-factor model. For their CAPM tests, GRS show that portfolios that rank near (far) each other have positively (negatively) correlated errors. That the $\hat{\mathbf{\Sigma}}$ matrix is not diagonal indicates that there are covariances that the risk-based model does not pickup. The difference-return setup has a near diagonal matrix.

The paper uses the generalized variance to compare different risk-based models and different test designs. Call the generalized variance of the dependent variables as $|\hat{\mathbf{\Sigma}}_{\mathbf{Y}}|$. Call the CAPM generalized variance $|\hat{\mathbf{\Sigma}}_{\mathbf{CAPM}}|$ and call the Fama and French 3-factor generalized variance $|\hat{\mathbf{\Sigma}}_{\mathbf{FF}}|$. Then the ratio $\frac{|\hat{\mathbf{\Sigma}}_{\mathbf{Y}}|-|\hat{\mathbf{\Sigma}}_{\mathbf{CAPM}}|}{|\hat{\mathbf{\Sigma}}_{\mathbf{Y}}|}$ is proportion of the generalized variance explained by CAPM and $\frac{|\hat{\mathbf{\Sigma}}_{\mathbf{y}}|-|\hat{\mathbf{\Sigma}}_{\mathbf{FF}}|}{|\hat{\mathbf{\Sigma}}_{\mathbf{Y}}|}$ is the proportion explained by the 3-factor model.[2] The generalized variance gives us to measure the additional precision from a change in the risk-based model. Tests also use Wilks λ in a MANOVA setup for this question. The MANOVA tests if the risk-based model shrinks the generalized variance of the errors. The MANOVA framework allows us to compare the standard and difference-return test designs. Tests show that the generalized variances of the difference-return tests is quite a bit smaller and do not change too much across the two risk-based models. The difference-return setup picks up a great deal, perhaps more, covariation than most complex risk-based models.

The remainder of the paper is organized as follows. Section 2 provides framework of the difference return, multivariate asset pricing model. Section 3 states the data and methods.

[2]These multivariate regression summaries have, of course, univariate counterparts in terms of R^2's.

Section 4 provide empirical tests. Section 5 concludes.

2 Setup

2.1 Notation

Let $\mathbf{y}_i$ be a $T \times 1$, vector of excess returns for dependent-variable portfolios $i = 1, 2, 3, ..., N$. Let $\mathbf{Y}$ be a $T \times N$ matrix of dependent-variable portfolios,

$$\mathbf{Y} = \begin{bmatrix} y_{1,1} & y_{2,1} & y_{3,1} & \cdots & y_{N,1} \\ y_{1,2} & y_{2,2} & y_{3,2} & \cdots & y_{N,2} \\ y_{1,1} & y_{2,3} & y_{3,3} & \cdots & y_{N,3} \\ . & . & . & . & \\ . & . & . & . & \\ . & . & . & . & \\ y_{1,T} & y_{2,T} & y_{3,T} & \cdots & y_{N,T} \end{bmatrix} = \begin{bmatrix} \mathbf{y}_1 & \mathbf{y}_2 & \mathbf{y}_3 & ... & \mathbf{y}_N \end{bmatrix}.$$

Let $\mathbf{X}$ be a $T \times (K + 1)$ matrix where the first column is vector of ones and the second through $K + 1$ columns are factors of the risk-based model. The multivariate regression is,

$$\mathbf{Y} = \mathbf{X}\hat{\mathbf{B}} + \hat{\mathbf{E}}, \tag{2}$$

with $\hat{\mathbf{B}}$ the performance measures and slopes on the factor and $\hat{\mathbf{E}}$ is a $T \times N$ matrix of errors. Assume that the errors are contemporaneously correlated, $cov(e_{i,t}, e_{j,t}) = \Sigma_{i,j}$; and $cov(e_{i,t}, e_{j,s}) = 0 \ \ \forall s \neq t$. $\hat{\mathbf{\Sigma}}$ has standard properties. MacKinlay and Richardson (1991) note that contemporaneous correlations may be hetereoskedastic, but this problem is small when T becomes large. The past literature typically does not make adjustments for the lack of homoskedasticity and for autocorrelation. This paper does not make adjustments to compare the results to those of other papers.

Assume standard multivariate theory. The vector of performance measures $\hat{\boldsymbol{\alpha}}$ is distributed as,

$$\hat{\boldsymbol{\alpha}} \sim \text{MVN}[\mathbf{0}, T^{-1}(1 + \hat{\Theta})^2\mathbf{\Sigma}], \tag{3}$$

where

$$\hat{\Theta} = \frac{\bar{r}_m}{s_p}$$

$$\bar{r}_m = \text{the sample mean of the market portfolio return } r_{m,t}, \text{ and}$$

$$s_p^2 = \text{estimate of the variance of the market portfolio } r_{m,t}.$$

$T(1+\hat{\Theta})^2\mathbf{\Sigma}$ follows a χ^2 distribution with N degrees of freedom. $\mathbf{\Sigma}$ is not known a priori and must be estimated. GRS use the W test statistic,

$$W = (1+\hat{\Theta}^2)^{-1}\hat{\boldsymbol{\alpha}}'\hat{\mathbf{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}. \tag{4}$$

GRS find that

$$\Gamma_{GRS} = [\frac{T}{T-K}][\frac{T-N-1}{N}]W \sim F_{N,T-N-1}, \tag{5}$$

where Γ_{GRS} has a central F distrbution. Tests use Γ_{GRS}.

2.2 The difference-return test design

The difference-return system is best explained by example. Rank all stocks by a characteristic and form three portfolios, $\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3$. Assume that there exists a risk-free rate and compute the excess stock returns for each portfolio $\mathbf{y}_i$. In the standard test, one runs a multivariate CAPM regression on the three portfolios,

$$\begin{aligned} \mathbf{y}_1 &= \alpha_1\mathbf{1}_T + \beta_1\mathbf{r}_m + \boldsymbol{\nu}_1 \\ \mathbf{y}_2 &= \alpha_2\mathbf{1}_T + \beta_2\mathbf{r}_m + \boldsymbol{\nu}_2 \\ \mathbf{y}_3 &= \alpha_3\mathbf{1}_T + \beta_3\mathbf{r}_m + \boldsymbol{\nu}_3, \end{aligned} \tag{6}$$

where $\mathbf{y}_i$ is a $T \times 1$ vector of observed excess stock returns, α_i is the performance measure, β_i is the slope on the market factor, $\mathbf{r}_m$ is a $T \times 1$ vector of excess returns of the market factor, and $\boldsymbol{\nu}_i$ is a $T \times 1$ vector of residuals. A second option is to compute the difference vectors, $\mathbf{d}_1 = \mathbf{y}_1 - \mathbf{y}_2$ and $\mathbf{d}_2 = \mathbf{y}_2 - \mathbf{y}_3$ and run the multivariate CAPM on the two difference vectors,

$$\begin{aligned} \mathbf{d}_1 &= a_1\mathbf{1}_T + b_1\mathbf{r}_m + \mathbf{u}_1 \\ \mathbf{d}_2 &= a_2\mathbf{1}_T + b_2\mathbf{r}_m + \mathbf{u}_2, \end{aligned} \tag{7}$$

where the regression setup is similar to the above 3-equation model. For this 2×2 difference-return setup, assume that the system is bivariate normal, such that $\boldsymbol{\Sigma}$,

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{1,2} \\ \sigma_{1,2} & \sigma_2^2 \end{bmatrix}, \boldsymbol{\Sigma}^{-1} = \frac{1}{\sigma_1^2\sigma_2^2 - (\sigma_{1,2})^2} \begin{bmatrix} \sigma_2^2 & -\sigma_{1,2} \\ -\sigma_{1,2} & \sigma_1^2 \end{bmatrix}, \tag{8}$$

where

$$\begin{aligned} \sigma_1^2 &= \frac{1}{T-2}\mathbf{u}_1'\mathbf{u}_1 = \frac{1}{T-2}[\boldsymbol{\nu}_1'\boldsymbol{\nu}_1 + \boldsymbol{\nu}_2'\boldsymbol{\nu}_2 - 2\boldsymbol{\nu}_1'\boldsymbol{\nu}_2] \\ \sigma_2^2 &= \frac{1}{T-2}\mathbf{u}_2'\mathbf{u}_2 = \frac{1}{T-2}[\boldsymbol{\nu}_2'\boldsymbol{\nu}_2 + \boldsymbol{\nu}_3'\boldsymbol{\nu}_3 - 2\boldsymbol{\nu}_2'\boldsymbol{\nu}_3] \\ \sigma_{1,2} &= \frac{1}{T-2}\mathbf{u}_1'\mathbf{u}_2 = \frac{1}{T-2}[\boldsymbol{\nu}_1'\boldsymbol{\nu}_2 - \boldsymbol{\nu}_2'\boldsymbol{\nu}_2 - \boldsymbol{\nu}_1'\boldsymbol{\nu}_3 + \boldsymbol{\nu}_2'\boldsymbol{\nu}_3]. \end{aligned} \tag{9}$$

Recall that GRS report cross-correlations of the errors. They find that portfolios that are rank close to each other have positively contemporaneously correlated errors and portfolios that rank far a way from each other have negatively correlated errors. If $\boldsymbol{\nu}_1'\boldsymbol{\nu}_2 > 0$, and $\boldsymbol{\nu}_2'\boldsymbol{\nu}_3 > 0$, then $\boldsymbol{\nu}_1'\boldsymbol{\nu}_2 - \boldsymbol{\nu}_2'\boldsymbol{\nu}_2 - \boldsymbol{\nu}_1'\boldsymbol{\nu}_3 + \boldsymbol{\nu}_2'\boldsymbol{\nu}_3 \approx 0$. Further, both $\boldsymbol{\nu}_1'\boldsymbol{\nu}_1 + \boldsymbol{\nu}_2'\boldsymbol{\nu}_2 - 2\boldsymbol{\nu}_1'\boldsymbol{\nu}_2$ and $\boldsymbol{\nu}_2'\boldsymbol{\nu}_2 + \boldsymbol{\nu}_3'\boldsymbol{\nu}_3 - 2\boldsymbol{\nu}_2'\boldsymbol{\nu}_3$ might be comparatively small. If so, the difference-return test $\hat{\boldsymbol{\Sigma}}$ is likely to be approximately diagonal.

2.3 Spectral decomposition of $\hat{\boldsymbol{\Sigma}}$

The eigenvalues of $\hat{\boldsymbol{\Sigma}}$ are important and help us understand the behavior of the system. The generalized variance is the sum of the eigenvalues. There is an additional benefit. For any quadratic form $\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}$,

$$\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha} = \frac{1}{\lambda_1}(\boldsymbol{\alpha}'\mathbf{e}_1)^2 + \frac{1}{\lambda_2}(\boldsymbol{\alpha}'\mathbf{e}_2)^2 + \frac{1}{\lambda_3}(\boldsymbol{\alpha}'\mathbf{e}_3)^2 + ... + \frac{1}{\lambda_N}(\boldsymbol{\alpha}'\mathbf{e}_N)^2, \tag{10}$$

where λ_i is the ith eigenvalue and $\mathbf{e}_i$ is its corresponding eigenvector. Tests multiply $\frac{1}{\lambda_i}(\boldsymbol{\alpha}'\mathbf{e}_i)^2$ term by a common multiplier, $[\frac{T}{T-K}][\frac{T-N-1}{N}](1+\hat{\Theta}^2)^{-1}$, for all the components so that the sum the multiplied components is Γ_{GRS} test statistic.

The spectral decomposition provides some insight into these tests. The eigenvalues are ranked in the amount of variation. If the $\hat{\alpha}$'s load primarily eigenvalues that have little variation, then the evidence suggests that the mispricing is connected with a nonrisk-based explanation or market inefficiency.

2.4 Power

GRS estimate the cross-sectional variances of the $\hat{\alpha}$'s and test for the relationship of the N portfolios and T time series observations. They find that power is best when N is about one third of T. MacKinlay (1987) makes two types of power tests. First, he ranks stocks by their $\hat{\beta}$'s, forms $\hat{\beta}$-ranked portfolios, and tests if the an error in the risk-free rate can be detected in CAPM. He finds power relatively poor. Second, he ranks stocks by market size and tests how easily one might detect a market-size factor as a supplementary risk-based factor to CAPM. He finds that the tests to detect missing factors have poor power.

GRS and MacKinlay (1987, 1995) use the following $\mathbf{\Sigma}$ pattern matrix. GRS use this matrix to evaluate the power properties of multivariate asset pricing tests. In the pattern matrix, ρ is the cross-correlation among the portfolios and is positive. For this matrix,

$$\mathbf{\Sigma} = (1-\rho)\sigma^2\mathbf{I}_N + \rho\sigma^2\mathbf{1}_N\mathbf{1}_N', \tag{11}$$

or,

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma^2 & \rho\sigma^2 & \rho\sigma^2 & ... & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 & \rho\sigma^2 & ... & \rho\sigma^2 \\ \rho\sigma^2 & \rho\sigma^2 & \sigma^2 & ... & \rho\sigma^2 \\ . & . & . & . & . \\ . & . & . & . & . \\ \rho\sigma^2 & \rho\sigma^2 & \rho\sigma^2 & . & \sigma^2 \end{bmatrix}. \tag{12}$$

The setup implies that

$$\mathbf{\Sigma}^{-1} = \frac{1}{(1-\rho)\sigma^2}[\mathbf{I}_N - \frac{\rho}{1+(N-1)\rho)}\mathbf{1}_N\mathbf{1}_N']. \tag{13}$$

The matrix is attractive because it is simple: its diagonal terms are positive constants and the off-diagonal terms are constants that are smaller than the diagonal terms. As $\rho \to 0$, the third term on the right-hand side drops out, and the equation simplifies. Their test design looks at the ratio of N to T, and they show that power is optimal when N is less than half of T.

The paper runs a horse race between the standard and difference return test designs. If such a pattern matrix holds, even approximately, then the difference return $\hat{\mathbf{\Sigma}}$ is approximately diagonal and takes the approximate form,

$$\mathbf{\Sigma} = \begin{bmatrix} 2(1-\rho)\sigma^2 & 0 & 0 & ... & 0 \\ 0 & 2(1-\rho)\sigma^2 & 0 & ... & 0 \\ 0 & 0 & 2(1-\rho)\sigma^2 & ... & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ 0 & 0 & 0 & . & 2(1-\rho)\sigma^2 \end{bmatrix}, \tag{14}$$

where ρ is positive and large and shrinks the diagonal terms.

For power tests, we use the empirically generated $\hat{\mathbf{\Sigma}}$'s. We assume that $\bar{\alpha} = 0$ and change the VAR(α) to test for power. Tests use two periods. In each period, tests compute CAPM and 3-factor tests for both the standard and difference-return test design.

2.5 Missing factor tests

Past papers focus on the use of principal components to generate a factor structure for pricing risk (See Connor and Korajczyk, for example.) The central idea is to find pervasive factors that explain the comovement of excess stock returns. One then breaks down their covariances into those explained by principal component factors and those that firm specific. The idea for principal components to explain risk is linked to the APT model of Ross (1977).

Assume first that CAPM holds for the first factor but there is also an additional factor is necessary to price returns. MacKinlay (1987, 1995) argues that the additional factor must (1) cause the α's to vanish, (2) shrink the generalized variance of the errors ($|\hat{\boldsymbol{Sigma}}|$), and (3) have a reasonably close Sharpe ratio to Sharpe ratio of the market portfolio. MacKinlay tests the power of the test to identify a missing factor. For a large second factor, where the difference between the high and low factor loadings is 12%, and the factor has a relatively low variance, power to identify the factor is quite good. When the premium is smaller and the variance of the factor larger, power falls off rapidly. He shows that if the missing factor has Sharpe ratio that approximates that of the market portfolio, the missing factor will be difficult to detect.

If the difference-return setup has more power, it is a better test for missing factors.

3 Empirical setup: methods and data

We use CRSP data for NYSE and AMEX stocks from 1926 to 2007. We two risk-based models, CAPM and the Fama and French 3-factor model. The data for SMB and HML are from Ken French's website. To sort stocks to form the dependent-variable portfolios, tests use market size and past performance rankings over the past six months. The study spans the period from July 1926 through December 2007. The study reports data for two periods, August 1926 through December 1966 and January 1967 through June 2007.

For market-size rankings, at June 30 of each year, we rank stocks by market size (price times shares outstanding) and form 10 portfolios. MV1 (MV10) is the smallest (largest) market-size decile. Portfolios are equally-weighted.

For 6-month past-performance rankings, each month we sum stock returns for the last six months and rank stocks and form 10 portfolios. P1 (P10) is the poorest (best) performing portfolio. Tests skip one month and invest. The momentum strategy is the triple (J,S,K) where J are the months of the ranking period, S are the months skipped between the rank and test periods, and K are the months of the holding period. Tests start at month t, rank stocks by past performance, group firms into 10 past-performance groups, skip 1 month between the ranking and test periods, and hold positions for the next 6 months. In this (6,1,6) strategy, the momentum return is the sum of six different long-short portfolios, 1/6 formed one month prior, 1/6 formed two months prior, and so forth. Tests repeat the process using month $t + 1$ as the portfolio formation date. Portfolios are equally-weighted.

Table I shows a comparison between the rebalanced and measurement-error-free returns. Blume and Stambaugh (1983), Boynton and Oppenheimer (2006) Conrad and Kaul (1993), Norman and Liu (2007), and Roll (1983) show that measurement errors may inflate portfolio returns. The measurement-error bias comes from bid-ask spread bounce that will typically inflate the portfolio returns of stocks that are small. Boynton and Oppenheimer and Norman and Liu show that small market-size stocks are prone to measurement errors and that measurement errors inflate the market-size premium and reduce the momentum premium. In the measurement-error-free return, the setup forms portfolio wealth relatives and computes monthly returns from adjacent monthly wealth relatives. Tests use the method of Boynton and Oppenheimer for making the measurement-error corrections.

Table I shows that in August 1926 through December 1966, the MV1 return is 0.491 too high and the MV10 return is 0.021 too high. The rebalanced returns *overstate* the MV1-MV10 premium by 0.470, or nearly by 6% per year. Similarly, the P1 rebalanced return is 0.290 per month too high and the P10 return is 0.021 per month too high. The rebalanced returns *understate* the P10-P1 momentum premium by 0.097 per month (or about 1% per year). In January 1967 through June 2007, the results are similar. The rebalanced returns *overstate* the MV10-MV1 premium by 0.556 (more than 6% per year) and *understate* the P10-P1, momentum premium by 0.154 (close to 2% per year). These results are consistent with those reported by Boynton and Oppenheimer and Norman and Liu. Because of these results, we use measurement-error-free returns for multivariate tests.

[Insert Table I about here.]

4 Results

4.1 Multivariate market-size sorts

There is a large literature that shows that small stocks have returns that are too high than that which we would predict by CAPM. Early papers include Banz (1981) and Reinganum (1982).[3] GRS and MacKinlay (1987, 1993) provide multivariate regression tests of the market-size premium.

Table II compares the slopes and performance measures of the two designs for both the CAPM and 3-factor risk-based models. Table II Panel A shows the standard parameter estimates for August 1926 to December 1966 period and the January 1967 through June 2007. Panel B shows the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ parameter estimates. Note that the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ parameters have smaller slopes on the factors and smaller performance measures.

[3]Fama and French (1992) show that β does not price the market-size premium, and they argue that market size and book-to-market are fundamental risk factors. Specific explanations for the small-firm premium include incomplete information (Barry and Brown, 1984), poorer economic fundamentals, higher levels of distress, and higher leverage (Chan and Chen, 1991), individual firm outliers in the smallest market-size decile (Knez and Ready, 1997), an underlying January effect (Keim, 1983, 1989; Blume and Stambaugh, 1983), transaction-cost frictions and higher bid-ask spreads (Stoll and Whaley, 1983 and Amihud and Mendelson, 1985), and a conditional asset pricing framework (Jagannathan and Wang, 1996).

[Insert Table II about here.]

Table III compares the cross-correlations of the errors. The matrices are shown for the test from August 1926 through June 2007. (Tests combine periods to save space; the subperiods are similar.) Panel A shows the cross-correlations of the standard test. The upper triangular matrix is the Fama and French 3-factor model and the lower matrix is CAPM. The Panel A results are similar to those that are reported in the past. The Panel A results show that the 3-factor model modestly shrinks correlations relative to those of CAPM.

Panel B results shows the cross-correlations of the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\alpha}}$ test. The Panel B results show that the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\hat{\boldsymbol{\alpha}}$ test has much smaller cross-correlations. For these tests, almost all of the paired cross-correlations do not reject. $\hat{\boldsymbol{\Sigma}}$ is approximately diagonal.

[Insert Table III about here.]

Each panel of Table IV shows two sets of is columns. The first set of five reports results for August 1926 through December 1966. The second set of five reports results for January 1967 through June 2007. The setup requires some discussion.

Consider the first set of five columns. The first three report the eigenvalues of the $\hat{\boldsymbol{\Sigma}}_{\mathbf{Y}}$, $\hat{\boldsymbol{\Sigma}}_{\mathbf{CAPM}}$, and $\hat{\boldsymbol{\Sigma}}_{\mathbf{FF}}$ matrices. To simplify reporting, the table mulitplies the eigenvalues by 1,000. The sum of the eigevalues is the generalized variance of the system. In a MANOVA, we also construct ratios of the generalized variances $|\hat{\boldsymbol{\Sigma}}_{\mathbf{Y}}|$, $|\hat{\boldsymbol{\Sigma}}_{\mathbf{CAPM}}|$, $|\hat{\boldsymbol{\Sigma}}_{\mathbf{FF}}|$, to test if the risk-based shrinks the generalized variance. The table reports the ratios of the $\frac{|\hat{\boldsymbol{\Sigma}}_{\mathbf{CAPM}}|}{|\hat{\boldsymbol{\Sigma}}_{\mathbf{Y}}|}$ and $\frac{|\hat{\boldsymbol{\Sigma}}_{\mathbf{FF}}|}{|\hat{\boldsymbol{\Sigma}}_{\mathbf{Y}}|}$ for tests of whether CAPM and the 3-factor model shrinks the variances.

In fourth and fifth columns, the table reports the GRS $\boldsymbol{\Gamma}_{CAPM}$ and $\boldsymbol{Gamma}_{FF}$ multivariate test statistics on joint performance. The spectral decomposition follows equation (11) but is multiplied by the $[\frac{T}{T-K}][\frac{T-N-1}{N}](1+\hat{\Theta}^2)^{-1}$ so that the sum of the components down equals the Γ_{GRS} test statistic for the multivariate joint test. To simplify reporting, the table mulitplies the eigenvalues by 1,000. The spectral decomposition in fourth and fifth columns is the loadings of the eigenvalueesystem on the $\hat{\boldsymbol{\alpha}}$'s. The loadings show which components have the highest contribution to abnormal performance and are useful for understanding better what is driving the detecting abnormal performance.

Table IV Panel A shows the results for the standard tests. Panel B shows the results for the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ tests. Across all tests, the GRS test stastistic is too small to reject the null of joint abnormal performance.

Table IV also shows that the generalized variance of the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ test is substantially smaller. The test results indicate that it is likely that the new test design will have better power.

[Insert Table IV about here.]

4.2 Momentum: rankings by 6-month, past performance

Jegadeesh and Titman (1993) were the first to show that recently ranked past winners outperform past losers.[4] Papers show that the momentum premium is not priced by standard asset pricing models, such as the CAPM or the 3-factor model.

Table V Panel A shows the parameter estimates for CAPM and the Fama and French 3-factor models. Panel A1 shows the results for the August 1926 to December 1966 and Panel A2 shows the results for the January 1967 through June 2007. The results are similar to those of Table 2. While the standard test has large loadings on the market, SMB, and HML factors, the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ test does not. Table V reports the GRS test statistics.

[Insert Table V about here.]

Table VI shows the cross-correlations of the error terms. In the standard tests, errors of closely ranked portfolios are positively correlated. In the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ shows very small cross-correlations. The $\hat{\boldsymbol{\Sigma}}$ is close to diagonal.

[Insert Table VI about here.]

[4]Jegadeesh and Titman (2001), Fama and French (1996), and Grundy and Martin (2001) show that momentum loses money in the U.S. pre-World War II era but earns money in the post-World War II era. Griffen, Ji, and Martin (2003) and Rouwenhorst (1998) find momentum in international markets. All these papers use univariate tests that go long in winners and short in losers.

Table VII shows the eigenvalues and spectral decomposition of the $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\hat{\boldsymbol{\alpha}}$ quadratic form. All test reject the null. The difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-\mathbf{1}}\hat{\boldsymbol{\alpha}}$ tests reject the null no joint differences among the α's. The standard tests reject the null of the α's being jointly different from zero.

[Insert Table VII about here.]

5 Conclusion

Multivariate regression tests in asset pricing ask if the performance measures are jointly different from zero. There are two related questions. First, we would like to know if there is a sort that is useful to generate unpriced premia for trading stocks. If the trade rule is useful, then an investor can sell the unprofitable position and buy a profitable one. Second, there may be common unpriced premia across all the portfolios. This paper concerns constructing a test that will help answer the question of making a test on whether the α's are jointly different from each other.

The proposed test computes the set of differences between two closely-ranked, dependent-variable portfolios and exploits the cross-correlated properties of portfolios that are ranked near each other. Tests find that these difference return portfolios generates a test with a smaller generalized variance and thus better power. The results will show that the market-size sort generates no abnormal in both types of risk-based tests and in both standard and difference-return designs. Momentum sorts generate premia in both risk-based models and test designs.

References

Amihud, Yakov, and Haim Mendelson. 1986. Asset pricing and the bid-asked spread. *Journal of Financial Economics*, 17 no. 2:223-249.

Banz, Rolf W. 1981. The relationship between return and market value of common stocks. *Journal of Financial Economics*, 9 no. 1:3-18.

Barry, Christopher B., and Stephen J. Brown. 1984. Differential information and the small firm effect. *Journal of Financial Economics*, 13 no. 2:283-294.

Blume, Marshall E., and Robert F. Stambaugh. 1982. Biases in computed returns: an application to the size effect. *Journal of Financial Economics*, 12 no. 3:387-404.

Boynton, Wentworth, and Henry. R. Oppenheimer. 2006. Anomalies in stock market pricing: problems in return measurements, *Journal of Business*, 79, no.5:2617-2631.

Chan, K. C., and Nai-Fu Chen. 1991. Structural and return characteristics of small and large firms. *Journal of Finance*, 46 no. 4:1467-1484.

Conrad, Jennifer, and Gautam Kaul. 1993. Long-term market overreaction or biases in computed returns? *Journal of Finance*, 48 no. 1:39-64.

Fama Eugene F., 1998. Market efficiency, long run returns, and behavioral finance. *Journal of Financial Economics*, 49, no. 3:283-306.

Fama, Eugene F., and Kenneth R. French. 1992. The cross-section of expected stock returns. *Journal of Finance*, 47 no. 2:427-466.

Fama, Eugene F., and Kenneth R. French. 1993. Common risk factors in the returns on common stocks and bonds. *Journal of Financial Economics*, 33 no. 1:3-56.

Gibbons, Michael R. 1982. Multivariate tests of financial models: a new approach. *Journal of Financial Economics*, 10 no. 1:3-27.

Gibbons, Michael R., Stephen Ross, and Jay Shanken. 1989. A test of efficiency of a given portfolio, *Econometrica*, 57 no. 5:1121-1152.

Grinblatt, Mark, and Sheridan Titman. 1985. Approximate facator structures: Interpretations and implications for empirical tests. *Journal of Finance*, 40 no. 5:1367-1373.

Grundy, Bruce D., and J. Spencer Martin. 2001. Understanding the nature of the risks and the source of the rewards to momentum investing. *Review of Financial Studies*, 14 no. 1:29-78.

Jagannathan, Ravi, and Zhenyu Wang. 1996. The conditional CAPM and the cross-section of expected returns. *Journal of Finance*, 51 no. 1:3-53.

Jegadeesh, Narasimhan, and Sheridan Titman. 1993. Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance*, 48 no. 1:65-92.

Jegadeesh, Narasimhan, and Sheridan Titman. 2001. Profitability of momentum strategies: an evaluation of alternative explanations. *Journal of Finance*, 56 no. 2:699-720.

Jobson, J.D., and Bob Korkie. 1982. Potential performance and tests of market efficiency. *Journal of Financial Economics*, 10 no. 4:433-466.

Johnson, Richard A., and Dean W. Wichern. 1992. *Applied Multivariate Statistical Analysis.* Prentice Hall, Englewood Cliffs, New Jersey.

Keim, Donald B. 1983. Size-related anomalies and stock return seasonality: further empirical evidence. *Journal of Financial Economics*, 12 no. 1:13-32.

Keim, Donald B. 1989. Trading patterns, bid-ask spreads, and estimated security returns: the case of common stocks at calendar turning points. *Journal of Financial Economics*, 25 no. 1:75-98.

Knez, Peter J., and Mark J. Ready. 1997. On the robustness of size and book-to-market in cross-sectional regressions. *Journal of Finance*, 52 no. 4:1355-1382.

Liu, Weimin and Norman Strong. 2006. Biases in decomposing holding period portfolio returns. Forthcoming, *Review of Financial Studies.*

MacKinlay, A. Craig. 1987. On multivariate tests of the CAPM. *Journal of Financial Economics*. 18 no. 2:341-372.

MacKinlay, A. Craig. 1995. Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics*, 38 no. 1:3-28.

MacKinlay, A. Craig, and Matthew P. Richardson, 1991. Using generalized method of moments to tests mean-variance efficiency. *Journal of Finance*, 46 no. 2:511-528.

Reinganum, Marc R. 1982. A direct test of Roll's conjecture on the firm size effect. *Journal of Finance*, 37 no. 1:27-35.

Roll, Richard. 1983. On computing mean returns and the small firm premium. *Journal of Financial Economics*, 12 no. 3:371-386.

Shanken, Jay. 1997. Statistical methods in tests of portfolio efficiency: a synthesis, in *Handbook of Statistics: Vol. 14: Statistical Methods in Finance*, G.S. Maddala and C.R. Rao (eds.), 693-711.

Stambaugh, Robert F. 1982. On the exclusion of assets from tests of the two-parameter model: a sensitivity analysis. *Journal of Financial Economics*, 10 no. 3:237-268.

Stoll, Hans R., and Robert E. Whaley. 1983. Transaction costs and the small firm effect. *Journal of Financial Economics*, 12 no. 1:57-80.

Table I: Market Size and 6-Month Past-Performance Portfolios: Rebalanced and Measurement-Error-Free Returns

Stock returns are from NYSE/AMEX from January 1926 through December 2007. Tests include all firms except for those that are certificates, American Depository Receipts (ADR's), units of beneficial interest, companies outside the U.S., Americus Trust certificates, closed-end funds, and real estate trusts. (Tests include stocks that have the share code classification of 10 or 11 on the CRSP file.) The setup ranks and forms portfolios in two ways. First, for market-size rankings, tests rank stocks by market capitalization (stock price times the number of shares) at June 30 each year, form 10 portfolios, and hold positions for July through June. MV1 (MV10) is the smallest (largest) market-size decile. Second, for momentum tests, tests rank stocks by 6-month past performance, skip one month, form portfolios, and hold stocks for the next 6 months. P1 (P10) is the worst (best) past-performing decile. All portfolios are equally-weighted. Panel A shows the monthly returns from August 1926 through December 1965, and Panel B shows the monthly returns from January 1967 through June 2007.

The table shows two types of returns. First, the table shows the rebalanced return, which are labelled as "REBAL." The rebalanced return is the cross-sectional average of the individual stock returns each month and is the standard in the literature. Second, the table uses a measurement-error-free, which are labelled as "MEF." Measurement-error bias comes from bid-ask spread bounce that will typically inflate the portfolio returns of stocks that are small. Thus, stocks that are small will have large spreads and rebalanced returns will have upward bias. The measurement-error-free metric uses the methodology from Boynton and Oppenheimer (2006). In the approach, the setup forms the time series of portfolio wealth relatives and computes monthly returns from the adjacent monthly wealth relatives.

Panel A: August 1926 - December 1966

	MV1	MV2	MV3	MV4	MV5	MV6	MV7	MV8	MV9	MV10
REBAL	1.624	0.817	0.803	0.779	0.817	0.755	0.787	0.740	0.679	0.522
MEF	1.132	0.808	0.831	0.793	0.814	0.762	0.770	0.715	0.652	0.501
Difference	0.491	0.010	-0.028	-0.014	0.003	-0.008	0.018	0.025	0.027	0.021

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
REBAL	1.417	1.214	1.159	1.157	1.140	1.121	1.130	1.211	1.297	1.543
MEF	1.128	1.071	1.046	1.066	1.044	1.038	1.046	1.126	1.205	1.350
Difference	0.290	0.143	0.113	0.091	0.096	0.082	0.083	0.084	0.092	0.193

Panel B: January 1967 - June 2007

	MV1	MV2	MV3	MV4	MV5	MV6	MV7	MV8	MV9	MV10
REBAL	2.359	1.581	1.338	1.227	1.123	1.131	1.022	0.930	0.930	0.754
MEF	1.798	1.331	1.214	1.183	1.079	1.118	1.046	0.922	0.933	0.749
Difference	0.561	0.251	0.124	0.045	0.044	0.013	-0.024	0.008	-0.003	0.005

	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10
REBAL	0.522	0.706	0.791	0.828	0.849	0.818	0.858	0.805	0.864	0.870
MEF	0.303	0.633	0.731	0.776	0.808	0.778	0.826	0.777	0.833	0.805
Difference	0.219	0.073	0.060	0.052	0.041	0.040	0.032	0.028	0.031	0.065

Table II: Market-size Portfolios: Parameter Estimates For CAPM and Fama and French 3-Factor Regressions

Stock returns are from NYSE/AMEX stocks from January 1926 through June 2007. Tests include all firms except for those that are certificates, American Depository Receipts (ADR's), units of beneficial interest, companies outside the U.S., Americus Trust certificates, closed-end funds, and real estate trusts. (Tests include stocks that have the share code classification of 10 or 11 on the CRSP file.) For market-size rankings, tests rank stocks by market capitalization (stock price times the number of shares) at June 30 each year, and form 10 market-size-ranked portfolios. Tests assume that investors holds positions from July through next June. MV1 (MV10) is the smallest (largest) market-size decile. Tests use unconditional CAPM and the Fama and French 3-factor model and use measurement-error-free returns.

The purpose of the table is to describe the parameter estimates for the intercepts and slopes for the two different designs. Panel A shows the standard multivariate regression with 10 portfolios. This test asks if the performance measures are jointly different from zero. Panel B shows the return difference test. The test asks if the performance measures are jointly different from each other. The return-difference test computes differences, MV1-MV2 (shown as MV1-2 and so forth), MV2-MV3, MV3-MV4,..., MV9-MV10, and regresses these differences on the chosen risk-based model. The joint test statistics on abnormal performance are reported in Table IV.

Panel A: Standard Multivariate Tests

	August 1926 through December 1966						January 1967 through June 2007					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	$\hat{h}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	$\hat{h}$
MV1	0.486	1.624	0.227	0.878	1.863	1.062	0.611	1.059	0.001	0.916	1.523	0.678
MV2	0.115	1.504	-0.061	1.006	1.390	0.634	0.277	1.078	-0.246	0.968	1.261	0.592
MV3	0.070	1.417	-0.064	1.042	1.095	0.451	0.305	1.069	-0.181	0.987	1.100	0.565
MV4	0.122	1.313	0.003	0.979	0.936	0.423	0.238	1.127	-0.210	1.071	0.947	0.536
MV5	0.032	1.296	-0.052	1.065	0.737	0.244	0.258	1.130	-0.166	1.075	0.902	0.507
MV6	0.097	1.263	0.032	1.074	0.470	0.270	0.212	1.119	-0.184	1.111	0.688	0.507
MV7	0.094	1.178	0.043	1.036	0.435	0.160	0.251	1.053	-0.101	1.084	0.479	0.481
MV8	0.019	1.118	-0.011	1.031	0.210	0.129	0.208	1.032	-0.076	1.088	0.276	0.411
MV9	0.047	1.096	0.027	1.029	0.070	0.147	0.154	1.011	-0.080	1.094	0.098	0.369
MV10	-0.036	0.972	-0.032	0.980	0.000	0.980	0.044	0.928	-0.045	1.022	0.000	1.022
Average	0.105	1.278	0.011	1.012	0.721	0.450	0.256	1.061	-0.143	1.056	0.639	0.555

Panel B: Return-Difference Tests

	August 1926 through December 1966						January 1967 through June 2007					
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	$\hat{h}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	$\hat{h}$
MV1-2	0.371	0.120	0.289	-0.128	0.473	0.428	0.334	-0.018	0.248	-0.052	0.261	0.085
MV2-3	0.046	0.088	0.003	-0.036	0.295	0.183	-0.027	0.009	-0.065	-0.019	0.162	0.027
MV3-4	-0.052	0.103	-0.067	0.063	0.159	0.028	0.067	-0.058	0.029	-0.084	0.152	0.029
MV4-5	0.089	0.017	0.055	-0.086	0.199	0.179	-0.020	-0.003	-0.044	-0.005	0.046	0.029
MV5-6	-0.065	0.033	-0.083	-0.009	0.267	-0.026	0.046	0.011	0.018	-0.036	0.214	0.000
MV6-7	0.003	0.085	-0.011	0.038	0.035	0.110	-0.040	0.066	-0.083	0.027	0.209	0.026
MV7-8	0.075	0.060	0.054	0.005	0.225	0.030	0.044	0.021	-0.025	-0.004	0.203	0.070
MV8-9	-0.029	0.022	-0.038	0.002	0.139	-0.017	0.053	0.021	0.004	-0.006	0.178	0.042
MV9-10	0.084	0.125	0.058	0.049	0.149	0.129	0.110	0.082	-0.035	0.072	0.279	0.180
Average	0.058	0.072	0.029	-0.011	0.216	0.116	0.063	0.015	0.005	-0.012	0.189	0.054

Table III: Portfolios From Market Size Rankings: Contemporaneous Cross-Correlation of Errors From CAPM and Fama and French 3-factor Regressions: August 1926 - June 2007

Stock returns are from NYSE/AMEX and stocks from January 1926 through June 2007. Tests include all firms except for those that are certificates, American Depository Receipts (ADR's), units of beneficial interest, companies outside the U.S., Americus Trust certificates, closed-end funds, and real estate trusts. (We include stocks that have the share code classification of 10 or 11 on the CRSP file.) Tests rank stocks by market capitalization (stock price times the number of shares) at June 30 each year, form 10 portfolios, and hold positions for July through June. MV1 (MV10) is the smallest (largest) market-size decile. Tests use the unconditional CAPM (lower triangular matrix) and the Fama and French 3-factor model (upper triangular matrix). Tests use measurement-error-free returns. (Tests with rebalanced returns are similar.)

There are two multivariate test designs. The first is the standard multivariate regression with 10 portfolios. The test asks if the performance measures are jointly different from zero. The second is the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ test that asks if the performance measures are jointly different from each other. This test computes differences, MV1-MV2, MV2-MV3, MV3-MV4, ..., MV9-MV10, and regresses these differences on the risk-based model.

Panel A: Standard Test Design: Crosscorrelations: CAPM lower triangular matrix; 3-factor upper triangular matrix

	MV1 (Small)	MV2	MV3	MV4	MV5	MV6	MV7	MV8	MV9	MV10 (Large)
MV1 (Small)		0.61	0.39	0.12	-0.06	-0.13	-0.17	-0.16	-0.17	-0.02
MV2	0.86		0.43	0.21	0.04	-0.04	-0.12	-0.11	-0.12	0.03
MV3	0.81	0.87		0.34	0.24	0.19	0.07	0.11	0.06	0.15
MV4	0.73	0.80	0.85		0.31	0.36	0.27	0.24	0.19	0.17
MV5	0.67	0.76	0.83	0.86		0.50	0.39	0.44	0.35	0.24
MV6	0.60	0.69	0.76	0.82	0.84		0.52	0.52	0.39	0.27
MV7	0.53	0.62	0.68	0.76	0.77	0.80		0.57	0.47	0.26
MV8	0.41	0.48	0.56	0.61	0.67	0.70	0.73		0.57	0.32
MV9	0.25	0.31	0.37	0.41	0.46	0.49	0.55	0.63		0.34
MV10 (Large)	-0.30	-0.31	0.15	-0.29	-0.26	-0.19	-0.18	-0.05	0.11	

Panel B: Return-Difference Test Design: Crosscorrelations: CAPM lower triangular matrix; 3-factor upper triangular matrix

	MV1-2	MV2-3	MV3-4	MV4-5	MV5-6	MV6-7	MV7-8	MV8-9	MV9-10
MV1-2		-0.09	0.17	0.12	0.10	-0.02	0.02	-0.05	-0.14
MV2-3	-0.02		-0.21	0.06	0.06	-0.03	0.04	-0.07	-0.13
MV3-4	0.20	-0.13		-0.35	0.13	0.03	-0.05	-0.01	-0.14
MV4-5	0.16	0.12	-0.27		-0.36	0.00	0.06	-0.05	-0.07
MV5-6	0.15	0.13	0.18	-0.25		-0.42	-0.03	-0.09	-0.09
MV6-7	0.10	0.07	0.11	0.09	-0.23		-0.42	0.05	-0.10
MV7-8	0.14	0.13	0.05	0.15	0.14	-0.18		-0.40	-0.03
MV8-9	0.07	0.04	0.07	0.03	0.09	0.17	-0.14		-0.31
MV9-10	0.10	0.08	0.05	0.10	0.18	0.11	0.24	-0.01	

Table IV: Market-Size Tests: Eigenvalues of the $\hat{\Sigma}$ and the Decomposition of the Gibbons, Ross, Shanken Test Statistic: August 1926 - June 2007

Tests rank stocks by the market size at June 30, form 10 portfolios, and hold positions for the next 12 months. MV1 (MV10) is the portfolio of smallest (largest) stocks. Tests use unconditional CAPM and the Fama and French 3-factor models and measurement-error-free returns. There are two periods, August 1926 through until December 1966 and January 1967 through June 2007. There are two multivariate test designs. The first is the standard multivariate regression with 10 portfolios. The test asks if the performance measures are jointly different from zero. The second is the difference-return test that evaluates if the α's are jointly different from each other. The difference-return test computes differences, MV1-MV2, MV2-MV3, MV3-MV4, ..., MV9-MV10, and regresses these differences on the risk-based model.

GRS use the W test statistic, $W = (1+\hat{\Theta}^2)^{-1}\hat{\boldsymbol{\alpha}}\ \hat{\boldsymbol{\Sigma}}^{\mathbf{-1}}\ \hat{\boldsymbol{\alpha}}$. GRS find that

$$\Gamma_{GRS} = [\frac{T}{T-K}][\frac{T-N-1}{N}]W \sim F_{N,T-N-1},$$

where Γ_{GRS} has a central F distrbution and this test statistic is used for tests.

One can use the spectral decomposition of any quadratic form $\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}$ is,

$$\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha} = \frac{1}{\lambda_1}(\boldsymbol{\alpha}'\mathbf{e}_1)^2 + \frac{1}{\lambda_2}(\boldsymbol{\alpha}'\mathbf{e}_2)^2 + \frac{1}{\lambda_3}(\boldsymbol{\alpha}'\mathbf{e}_3)^2 + ... + \frac{1}{\lambda_N}(\boldsymbol{\alpha}'\mathbf{e}_N)^2.$$

where λ_i is the ith eigenvalue and $\mathbf{e}_i$ is its corresponding eigenvector. We will use this decomposition in empirical tests. One can multiply $\frac{1}{\lambda_i}(\boldsymbol{\alpha}'\mathbf{e}_i)^2$ term by a common multiplier, $[\frac{T}{T-K}][\frac{T-N-1}{N}](1+\hat{\Theta}^2)^{-1}$, for all the components. Then we sum the multiplied components for the Γ_{GRS} test statistic.

Each panel shows the eigenvalues and the decomposition of the quadratic form. (The eigenvalues are all multiplied by 1,000 to reduce clutter from zeros.) The table is organized as follows. The largest eigenvalues are at the top of the table. The sum of the eigenvalues is the generalized variance. The smaller the generalized variance, the greater the likelihood of improved power. The GRS test statistic is shown as well as the loadings on each of the eigenvalues. The sum of these loadings equals the GRS W test statistic.

None of the GRS W statistics are not large enough to reject the null at $p < .05$.

Panel A: Standard Test Design: 10 Market Size Portfolios

	August 1926 to December 1966					January 1967 to June 2007				
E	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}
1	75.574	14.259	2.025	0.090	0.107	30.750	9.938	2.430	0.415	0.115
2	4.288	1.142	0.370	0.113	0.171	3.232	1.551	1.170	0.018	0.578
3	0.651	0.370	0.255	0.244	0.150	0.520	0.344	0.259	0.747	0.674
4	0.249	0.219	0.215	0.040	0.012	0.224	0.204	0.172	0.116	0.001
5	0.210	0.148	0.145	0.023	0.016	0.148	0.126	0.100	0.032	0.190
6	0.124	0.124	0.123	0.001	0.001	0.098	0.095	0.084	0.039	0.078
7	0.121	0.100	0.092	0.003	0.042	0.086	0.068	0.065	0.009	0.000
8	0.092	0.090	0.069	0.103	0.072	0.057	0.056	0.056	0.202	0.069
9	0.078	0.066	0.064	0.095	0.233	0.053	0.048	0.041	0.073	0.137
10	0.062	0.033	0.031	0.267	0.163	0.033	0.032	0.028	0.006	0.162
GV	81.449	16.551	3.388	0.978	0.967	35.200	12.463	4.405	1.656	2.004
		79.7%	95.8%				64.6%	87.5%		
GRS				0.978	0.967				1.656	2.004

Panel B: Return-Difference Test Design: Market Size Portfolios

	August 1926 to December 1966					January 1967 to June 2007				
E	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}
1	2.376	2.270	1.587	0.338	0.273	0.890	0.890	0.772	0.685	0.404
2	0.799	0.797	0.726	0.007	0.045	0.405	0.403	0.376	0.074	0.090
3	0.498	0.441	0.391	0.134	0.080	0.335	0.333	0.324	0.005	0.004
4	0.356	0.319	0.295	0.004	0.005	0.250	0.231	0.217	0.142	0.060
5	0.265	0.254	0.206	0.000	0.025	0.215	0.212	0.170	0.005	0.017
6	0.217	0.210	0.200	0.159	0.148	0.174	0.170	0.121	0.000	0.002
7	0.189	0.184	0.118	0.001	0.000	0.121	0.121	0.083	0.080	0.156
8	0.096	0.092	0.081	0.016	0.263	0.086	0.085	0.073	0.492	0.078
9	0.087	0.080	0.010	0.229	0.005	0.073	0.072	0.026	0.001	0.949
GV	4.884	4.647	3.614			2.549	2.517	2.163		
		4.8%	26.0%				1.3%	15.2%		
GRS				0.887	0.844				1.484	1.759

Table IV: Momentum Portfolios: Parameter Estimates for Performance and Slopes

Stock returns are from NYSE/AMEX stocks from January 1926 through June 2007. Tests include all firms except for those that are certificates, American Depository Receipts (ADR's), units of beneficial interest, companies outside the U.S., Americus Trust certificates, closed-end funds, and real estate trusts. (Tests include stocks that have the share code classification of 10 or 11 on the CRSP file.) Tests rank stocks by 6-month past performance (sum of the last 6 months of stock returns) each month, and form 10 portfolios. The test assumes that the investor will skip a month between the ranking and test periods and hold positions for the next 6 months. P1 (P10) is the portfolio of past losers (winners). Tests use unconditional CAPM and the Fama and French 3-factor model and use measurement-error-free returns.

The purpose of the table is to describe the parameter estimates for the intercepts and slopes for the two different designs. The first is the standard multivariate regression with 10 portfolios, which tests if the performance measures are jointly different from zero. The second is the modified test proposed to test if the performance measures are jointly different from each other. The modified test computes differences, P1-P2 (shown as P1-2), P2-P3, P3-P4, ..., P9-P10, and regresses these differences on the risk-based model.

The joint test statistics on abnormal performance are reported in Table VIII.

Panel A: Standard Multivariate Tests

August 1926 through December 1966

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	h
P1	-0.096	1.515	-0.255	1.040	1.280	0.627
P2	-0.078	1.422	-0.193	1.095	0.851	0.450
P3	-0.025	1.327	-0.123	1.051	0.730	0.373
P4	0.045	1.264	-0.035	1.035	0.612	0.307
P5	0.048	1.233	-0.030	1.015	0.609	0.279
P6	0.079	1.188	0.012	0.994	0.566	0.232
P7	0.097	1.174	0.033	0.980	0.542	0.246
P8	0.177	1.175	0.112	0.986	0.578	0.213
P9	0.225	1.214	0.150	0.995	0.673	0.245
P10	0.289	1.313	0.184	0.999	0.002	0.999
Average	0.076	1.283	-0.014	1.049	0.611	0.567

January 1967 through June 2007

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	$\hat{h}$
P1	-0.308	1.272	-0.761	1.170	1.161	0.555
P2	0.092	1.127	-0.314	1.100	0.815	0.549
P3	0.223	1.056	-0.161	1.058	0.673	0.541
P4	0.297	0.999	-0.064	1.008	0.603	0.509
P5	0.336	0.981	-0.028	1.002	0.572	0.524
P6	0.311	0.971	-0.036	0.994	0.535	0.498
P7	0.353	0.985	0.015	0.994	0.563	0.475
P8	0.291	1.012	-0.039	1.022	0.548	0.467
P9	0.319	1.070	-0.018	1.056	0.645	0.459
P10	0.236	1.183	-0.113	1.090	-0.001	1.090
Average	0.215	1.066	-0.152	1.049	0.611	0.567

Panel B: Return-Difference Tests

August 1926 through December 1966

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	h
P1-P2	-0.018	0.092	-0.071	-0.055	0.431	0.176
P2-P3	-0.052	0.096	-0.070	0.045	0.121	0.076
P3-P4	-0.070	0.063	-0.086	0.016	0.119	0.066
P4-P5	-0.003	0.031	-0.007	0.020	0.006	0.026
P5-P6	-0.031	0.046	-0.039	0.020	0.042	0.048
P6-P7	-0.019	0.013	-0.019	0.014	0.025	-0.015
P7-P8	-0.079	-0.001	-0.080	-0.006	-0.037	0.034
P8-P9	-0.048	-0.038	-0.037	-0.008	-0.095	-0.032
P9-P10	-0.064	-0.099	-0.030	-0.004	-0.290	-0.108
Average	-0.038	0.020	-0.049	0.005	0.036	0.030

January 1967 through June 2007

	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\alpha}$	$\hat{\beta}$	$\hat{s}$	$\hat{h}$
P1-P2	-0.400	0.145	-0.444	0.071	0.345	0.006
P2-P3	-0.132	0.071	-0.154	0.042	0.143	0.009
P3-P4	-0.073	0.057	-0.100	0.051	0.069	0.034
P4-P5	-0.040	0.018	-0.035	0.007	0.030	-0.015
P5-P6	0.025	0.010	0.005	0.009	0.038	0.027
P6-P7	-0.042	-0.013	-0.052	0.000	-0.029	0.024
P7-P8	0.062	-0.027	0.057	-0.029	0.017	0.006
P8-P9	-0.029	-0.058	-0.021	-0.034	-0.098	0.007
P9-P10	0.083	-0.114	0.093	-0.036	-0.297	0.043
Average	-0.054	0.009	-0.072	0.009	0.024	0.016

Table VI: Momentum: Contemporaneous Cross-Correlation of Errors: August 1926 - June 2007

Stock returns are from NYSE/AMEX stocks from January 1926 through June 2007. Tests include all firms except for those that are certificates, American Depository Receipts (ADR's), units of beneficial interest, companies outside the U.S., Americus Trust certificates, closed-end funds, and real estate trusts. (Tests include stocks that have the share code classification of 10 or 11 on the CRSP file.) For 6-month, past-performance rankings, tests rank stocks by the sum of their 6 month past returns, form 10 past-performance-ranked portfolios, skip one month and hold positions for the next 6 months. P1 (P10) is the portfolio of past losers (winners). Tests use unconditional Capital Asset Pricing Theory and the Fama and French 3-factor models. Tests use measurement-error-free returns.

There are two multivariate test designs. The first is the standard multivariate regression with 10 portfolios. The test asks if the performance measures are jointly different from zero. The second is the difference-return $\hat{\boldsymbol{\alpha}}\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$ test that asks if the performance measures are jointly different from each other. This test computes differences, P1-P2, P2-P3, P3-P4, ..., P9-P10, and regresses these differences on the risk-based model.

Panel A: Standard Test Design: Crosscorrelations: CAPM lower triangular matrix; 3-factor upper triangular matrix

	P1 (Losers)	P2	P3	P4	P5	P6	P7	P8	P9	P10 (Winners)
P1 (Losers)		0.69	0.56	0.44	0.35	0.28	0.24	0.19	0.21	0.43
P2	0.91		0.79	0.71	0.64	0.53	0.46	0.34	0.25	0.24
P3	0.88	0.96		0.79	0.75	0.66	0.57	0.42	0.30	0.18
P4	0.85	0.94	0.96		0.76	0.73	0.64	0.53	0.36	0.18
P5	0.83	0.92	0.94	0.95		0.76	0.69	0.57	0.43	0.20
P6	0.81	0.89	0.92	0.94	0.95		0.76	0.69	0.52	0.26
P7	0.80	0.88	0.90	0.91	0.93	0.94		0.71	0.60	0.35
P8	0.80	0.86	0.87	0.89	0.91	0.93	0.94		0.73	0.47
P9	0.80	0.84	0.84	0.85	0.87	0.89	0.92	0.95		0.65
P10 (Winners)	0.84	0.81	0.79	0.80	0.80	0.82	0.85	0.88	0.91	

Panel B: Return-Difference Test Design: Crosscorrelations: CAPM lower triangular matrix; 3-factor upper triangular matrix

	P1-2	P2-3	P3-4	P4-5	P5-6	P6-7	P7-8	P8-9	P9-10
P1-2		0.19	0.18	0.10	0.03	0.02	-0.06	-0.15	-0.50
P2-3	0.33		-0.14	0.14	0.10	0.03	-0.04	-0.07	-0.26
P3-4	0.28	-0.01		-0.33	0.20	0.00	0.13	-0.02	-0.07
P4-5	0.15	0.19	-0.27		-0.33	0.09	-0.01	0.03	-0.05
P5-6	0.10	0.16	0.22	-0.30		-0.30	0.13	0.07	0.04
P6-7	-0.01	0.02	-0.01	0.08	-0.30		-0.34	0.17	0.10
P7-8	-0.09	-0.06	0.09	-0.02	0.13	-0.33		-0.19	0.10
P8-9	-0.27	-0.19	-0.09	-0.03	-0.01	0.19	-0.17		0.09
P9-10	-0.59	-0.40	-0.21	-0.12	-0.07	0.12	0.11	0.22	

Table VII: Momentum: Eigenvalues and Gibbons, Ross, Shanken Test Statistics: August 1926 - June 2007

Tests rank stocks by the sum of the their returns in the last 6 months, form 10 portfolios, skip one month and hold positions for the next 6 months. P1 (P10) is the portfolio of past losers (winners). Tests use unconditional CAPM and the Fama and French 3-factor models and measurement-error-free returns. There are two periods for tests. The first is from August 1926 through until December 1966. The second is from January 1967 through June 2007. There are two multivariate test designs. The first is the standard multivariate regression with 10 portfolios. The test asks if the performance measures are jointly different from zero. The second is the difference-return test that evaluates if the α's are jointly different from each other. The modified test computes differences, P1-P2, P2-P3, P3-P4, ..., P9-P10, and regresses these differences on the risk-based model.

GRS use the W test statistic, $W = (1+\hat{\Theta}^2)^{-1}\hat{\boldsymbol{\alpha}}'\hat{\boldsymbol{\Sigma}}^{-1}\hat{\boldsymbol{\alpha}}$. GRS find that

$$\Gamma_{GRS} = [\frac{T}{T-K}][\frac{T-N-1}{N}]W \sim F_{N,T-N-1},$$

where Γ_{GRS} has a central F distrbution and this test statistic is used for tests.

One can use the spectral decomposition of any quadratic form $\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha}$ is,

$$\boldsymbol{\alpha}'\boldsymbol{\Sigma}^{-1}\boldsymbol{\alpha} = \frac{1}{\lambda_1}(\boldsymbol{\alpha}'\mathbf{e}_1)^2 + \frac{1}{\lambda_2}(\boldsymbol{\alpha}'\mathbf{e}_2)^2 + \frac{1}{\lambda_3}(\boldsymbol{\alpha}'\mathbf{e}_3)^2 + ... + \frac{1}{\lambda_N}(\boldsymbol{\alpha}'\mathbf{e}_N)^2.$$

where λ_i is the ith eigenvalue and $\mathbf{e}_i$ is its corresponding eigenvector. We will use this decomposition in empirical tests. One can multiply $\frac{1}{\lambda_i}(\boldsymbol{\alpha}'\mathbf{e}_i)^2$ term by a common multiplier, $[\frac{T}{T-K}][\frac{T-N-1}{N}](1+\hat{\Theta}^2)^{-1}$, for all the components. Then we sum the multiplied components for the Γ_{GRS} test statistic.

Each panel shows the eigenvalues and the decomposition of the quadratic form. (The eigenvalues are all multiplied by 1,000 to reduce clutter from zeros.) The table is organized as follows. The largest eigenvalues are at the top of the table. The sum of the eigenvalues is the generalized variance. The smaller the generalized variance, the greater the likelihood of improved power. The GRS test statistic is shown as well as the loadings on each of the eigenvalues. The sum of these loadings equals the GRS W test statistic.

All GRS are large enough to reject the null at $p < .05$.

Panel A: Standard Test Design: 10 Momentum Portfolios

E	August 1926 to December 1966					January 1967 to June 2007				
	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}
1	73.520	9.561	0.733	0.017	0.378	30.617	7.807	1.991	0.194	1.092
2	0.795	0.414	0.403	1.057	0.865	0.683	0.561	0.490	4.982	2.578
3	0.406	0.330	0.245	1.329	1.015	0.303	0.303	0.237	0.001	0.072
4	0.072	0.069	0.068	0.097	0.100	0.034	0.033	0.032	0.246	0.184
5	0.051	0.051	0.050	0.057	0.049	0.020	0.020	0.019	2.432	1.768
6	0.043	0.043	0.043	0.007	0.017	0.014	0.014	0.014	0.028	0.023
7	0.029	0.029	0.028	0.059	0.052	0.013	0.013	0.013	0.003	0.025
8	0.027	0.027	0.027	0.041	0.073	0.010	0.010	0.010	0.744	0.667
9	0.022	0.022	0.022	0.070	0.042	0.010	0.009	0.009	0.343	0.291
10	0.019	0.019	0.019	0.242	0.197	0.009	0.009	0.009	0.445	0.378
GV	74.984	10.565	1.637			31.712	8.779	2.824		
		85.9%	97.8%				72.3%	91.1%		
GRS				2.976	2.788				9.417	7.078

Panel B: Return-Difference Test Design: Momentum Portfolios

E	August 1926 to December 1966					January 1967 to June 2007				
	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}	Σ_{Y}	Σ_{CAPM}	Σ_{FF}	Γ_{CAPM}	Γ_{FF}
1	0.981	0.883	0.478	0.001	0.016	0.676	0.589	0.363	1.441	3.222
2	0.215	0.210	0.196	0.062	0.092	0.084	0.082	0.070	1.503	0.883
3	0.134	0.106	0.097	0.006	0.027	0.050	0.049	0.049	0.453	0.282
4	0.090	0.090	0.087	0.087	0.077	0.040	0.038	0.036	0.210	0.528
5	0.082	0.080	0.075	0.004	0.009	0.037	0.037	0.036	0.552	0.436
6	0.071	0.071	0.068	0.505	0.493	0.030	0.027	0.026	0.484	0.930
7	0.053	0.046	0.038	1.161	1.984	0.021	0.021	0.021	0.006	0.001
8	0.029	0.028	0.026	0.117	0.028	0.016	0.015	0.014	0.000	0.030
9	0.022	0.022	0.022	0.062	0.123	0.012	0.012	0.012	2.104	1.523
GV	1.676	1.536	1.086			0.965	0.870	0.626		
		8.4%	35.2%				9.8%	35.1%		
GRS				2.006	2.850				6.753	7.835