Portfolio Correlation and the Power of Portfolio Efficiency Tests*

Debrah Meloso[†] Peter Bossaerts[‡]

January, 2008

Abstract

We propose a parametric family of tests of the mean-variance efficiency of a portfolio in a market with a risk-free asset. All tests in the family compare the mean-variance ratio of the tested portfolio (the *benchmark* portfolio) with the same ratio for a different portfolio, called the *reference* portfolio. The Gibbons-Ross-Shanken test belongs to this family, and the reference in this case is the ex-post tangency portfolio of the market. We show that the power of a test in our proposed family depends on the correlation between the benchmark and the reference portfolio. This correlation, and thus, the power of the test, can be manipulated by changing the value of the parameter that spans the family. In particular, for a given sample, a power maximizing test can be easily found in the family we propose. This power-maximizing test will generically not be the Gibbons-Ross-Shanken test.

Keywords: Mean-variance efficiency, GRS-test, Sharpe ratio, Statistical power.
JEL Classification Numbers: G12, C13

1 Introduction

The mean and the variance of a portfolio are the most commonly used and reported measures of its goodness. Mean-variance efficiency is a minimal criterion to sort portfolios along these two dimensions of performance, and is therefore widely used and reported. On the other hand, the mean-variance efficiency of the *market* portfolio is one of the main

*This paper has benefitted from comments received from Peter Phillips and seminar audiences at the University of Auckland, and the School of Management at Yale University.
[†]ESC-Rennes School of Business. Corresponding author, email: debrah.meloso@esc-rennes.com
[‡]California Institute of Technology and CEPR.

results of the CAPM model, adding to the reasons why the measurement of mean-variance efficiency is relevant.[1]

If a portfolio is mean-variance efficient, the mean returns of all assets can be represented as coming from two sources, one of them being the mean-variance efficient portfolio, and the other being an asset with fixed returns. In other words, mean excess returns (excess over the return of an asset with fixed return) of all assets are a linear transformation of the mean excess return of a mean-variance efficient portfolio. If the distribution of returns of all assets is known, the above result can be readily used to establish what portfolios are mean-variance efficient. In reality, the distribution of returns is unknown and must be recovered from a sample of realizations. The statistical equivalent to verifying the linear transformation result is to establish whether the *Jensen* alphas are statistically different from zero. This question spans a class of tests used to establish mean-variance efficiency of an arbitrary portfolio, to test the CAPM if the market portfolio is used, and to test other, non-equilibrium, asset pricing models, like the multi-factor models. This is why much of what is claimed about the empirical success or failure of asset pricing models, as well as about portfolio performance in general, depends on the statistical properties of this class of tests.

We propose a family of tests that lies in this class. It is based on the GRS test proposed in Gibbons, Ross, and Shanken [1989], and comprises this test as a special case. Under the assumption that a risk-free asset exists, the GRS test is nowadays the standard test of mean-variance efficiency, because of its simple way of dealing with the twofold statistical problem of multiple hypothesis testing and serial correlation. It has, however, been subjected to the criticism of having low statistical power.[2] This is the issue we deal with. On one hand, we point at the source of the power properties of the family that we propose (including the GRS test), and the Gibbons test (Gibbons [1982]).[3] On the other hand, we reduce the problem of finding a powerful test - within the family we propose - to that of changing a single parameter.

The idea behind the possibility of improving upon the power of the GRS test is very simple. Suppose a statistician is faced with the task of finding out whether the means of two random variables, X and Y, are equal. To do this, she is endowed with two tools. One is a

[1] Mean-variance efficiency of the market portfolio can also be derived as an "approximate" result in equilibrium pricing models with assumptions on preferences or return distributions that differ from those in the CAPM.

[2] The power of a test is the probability of rejecting the null hypothesis when it is false.

[3] The Gibbons test is the equivalent of the GRS test in the absence of a risk-free asset. If used in the presence of a risk-free asset, it is in principle a weaker test than the GRS. Nonetheless, the Gibbons test has in occasions rejected the null hypothesis where the GRS hasn't. This fact lies at the heart of the claim that the GRS has low power.

finite sample taken from the true joint distribution of X and Y. The other tool is the power of choosing the correlation of X and Y. It is clear that, given such power, she will choose maximal correlation, since this minimizes the variance of the statistical inference she can make from the sample. Statisticians rarely have this power. However, the econometrician that is trying to assess the mean-variance efficiency of a portfolio implicitly has this power. The family of test statistics that we propose is based on the comparison between two portfolios (like the comparison between X and Y that the statistician had to make). The GRS test fixes this comparison to be between the sample mean-variance efficient portfolio and the portfolio whose mean variance efficiency is being tested. We claim that the latter need not be compared to the sample mean-variance efficient portfolio in order to asses its efficiency. In fact, the family of tests we propose is given by a parameter identifying other candidates for comparison. Picking a different candidate is equivalent to choosing a different covariance for the random variables being compared. This is how statistical power can be increased.

Of course, the econometrician does not know the true covariance of portfolios, that can lead to improved power. An estimate must be used instead. We use this estimate to find the power maximizing test within the family we propose, for a sample of CRSP data on portfolio returns. We show that power improvements over the GRS test can be very significant.

In the following section we introduce notation and define mean-variance efficiency. In section 3 we propose a family of mean-variance efficiency tests and give its statistical properties. In Section 4 we develop the intuition that tests of mean-variance efficiency are based on the comparison between two portfolios, and give the result that the power of the test depends on the true correlation between these portfolios. We analyze the power of the proposed family of tests in Section 5. In Section 6 we give numerical examples of applications of our proposed power improving tests. We conclude in Section 7, suggesting extensions of this work, and a possible answer to the puzzle mentioned in footnote 3.

2 Problem Statement

Throughout this paper we will use *efficiency* to refer to mean-variance efficiency of a portfolio. We proceed to define this latter notion, after giving the necessary setup and notation.

Let $\{\bar{R}, \Omega\}$ denote a market. For our purpose a market is fully defined by mean and covariance of asset returns. Risky assets have random returns denoted $R = (R_1, \ldots, R_N)$.

$\bar{R} = \left(\bar{R}_1, \ldots, \bar{R}_N\right)$ is the vector of mean returns of risky assets, and Ω is the $N \times N$ matrix of asset return covariances. We will mostly consider markets where there is a risk-free asset, with fixed, known return R_F. For clarity, we will use $\{\bar{R}, \Omega\}$ whenever referring to a market with no risk-free asset, and $\{\bar{R}, \Omega, R_F\}$ whenever referring to a market with a risk-free asset.

A portfolio, p, is a combination of assets. Let $w_p = (w_{p1}, \ldots, w_{pN})$ denote the relative holdings of each risky asset that makes up portfolio p. Each weight is the fraction of total wealth invested in the corresponding asset. The residual, $(1 - w_p'\mathbf{1})$, gives the relative holding of the risk-free security. The mean return of portfolio p is $\bar{R}_p = w_p'\bar{R} + (1 - w_p'\mathbf{1})R_F$, and its variance is $\sigma_p^2 = w_p'\Omega w_p$. The covariance between portfolio p and all risky assets is the vector $(\sigma_{p1}, \ldots, \sigma_{pN}) = w_p'\Omega$, and the covariance of two portfolios p and q is $\sigma_{pq} = w_p'\Omega w_q$.

The Sharpe ratio of portfolio p is defined to be

$$SR^p = \frac{\bar{R}_p - R_F}{\sigma_p}. \tag{1}$$

The Sharpe ratio of a portfolio is its measure of efficiency. Given $\{\bar{R}, \Omega, R_F\}$, a portfolio is *efficient* if it has maximal Sharpe ratio. Alternatively, we say that *portfolio e is efficient* if w_e solves the following program:

$$\begin{aligned} &\min_w w'\Omega w \qquad (2) \\ &\text{s.t.} \quad \bar{R}_e = w'\bar{R} + (1 - w'\mathbf{1})R_F \end{aligned}$$

for a target mean return $\bar{R}_e$.

It is clear from (2) that there are many efficient portfolios, depending on the chosen target mean return, $\bar{R}_e$. All efficient portfolios in $\{\bar{R}, \Omega, R_F\}$ have the same Sharpe ratio, which clearly means that they are located on a line in mean-variance space. In other words, if program (2) is solved for all R_e, the solution is a linear relation between the mean and standard deviation of efficient portfolios, with slope equal to the maximal Sharpe ratio. Specifically, the solutions to program (2) define the following *efficient frontier* for market $\{\bar{R}, \Omega, R_F\}$:

$$\bar{R}_e = R_F + \left(\sqrt{(\bar{R} - R_F\mathbf{1})'\Omega^{-1}(\bar{R} - R_F\mathbf{1})}\right)\sigma_e, \tag{3}$$

and the maximal Sharpe ratio in market $\{\bar{R}, \Omega, R_F\}$, is the slope in the above expression,

$$SR^* = SR^e = \frac{\bar{R}_e - R_F}{\sigma_e} = \sqrt{\left(\bar{R} - R_F\mathbf{1}\right)' \Omega^{-1} \left(\bar{R} - R_F\mathbf{1}\right)} \tag{4}$$

The efficient frontier in market $\{\bar{R}, \Omega, R_F\}$ is affine because of the presence of the risk-free asset. One can linearly trade off mean return and standard deviation by changing the relative holding of the risk-free asset with respect to a fixed portfolio of risky assets. This fixed portfolio of risky assets is the *tangency portfolio* corresponding to R_F, which will be denoted by t, and has weights $w_t = \frac{\Omega^{-1}(\bar{R}-R_F\mathbf{1})}{\mathbf{1}'\Omega^{-1}(\bar{R}-R_F\mathbf{1})}$.

The tangency portfolio for market $\{\bar{R}, \Omega, R_F\}$, is efficient in this market as well as in market $\{\bar{R}, \Omega\}$, where there is no risk-free asset. It lies at the tangency of the efficient frontier for market $\{\bar{R}, \Omega, R_F\}$ and that of market $\{\bar{R}, \Omega\}$. The latter is not a line in mean-standard deviation space. It is a parabola that lies within the efficient frontier for the market with a risk-free asset, since more combinations of assets can be achieved in the larger market, $\{\bar{R}, \Omega, R_F\}$. Figure 1 shows an example of efficient frontier for a market without risk-free asset, the efficient frontier once a risk-free asset is added, and the tangency portfolio for $\{\bar{R}, \Omega, R_F\}$.

An important property of an efficient portfolio is that, for all assets in the market, the excess return (difference of the asset's return and the return of the risk-free asset) of an asset is a linear transformation of the excess return of the efficient portfolio. This property lies at the heart of the family of tests that we will propose, as well as the GRS test of Gibbons, Ross, and Shanken [1989]. Thus, for $n \in \{1, \ldots, N\}$ it holds that

$$\bar{R}_n - R_F = \frac{\sigma_{en}}{\sigma_e^2}(\bar{R}_e - R_F). \tag{5}$$

For the sake of completeness, we include one derivation of the above expression. Define a portfolio ϕ with the following mean and variance of returns:

$$\begin{aligned} \bar{R}_\phi &= \phi\bar{R}_e + (1-\phi)\,\bar{R}_p \\ \sigma_\phi &= \left[\phi^2\sigma_e^2 + (1-\phi)^2\,\sigma_p^2 + 2\phi\,(1-\phi)\,\sigma_{ep}\right]^{\frac{1}{2}}, \end{aligned}$$

where p is an arbitrary portfolio. The Sharpe ratio of the efficient portfolio can be found

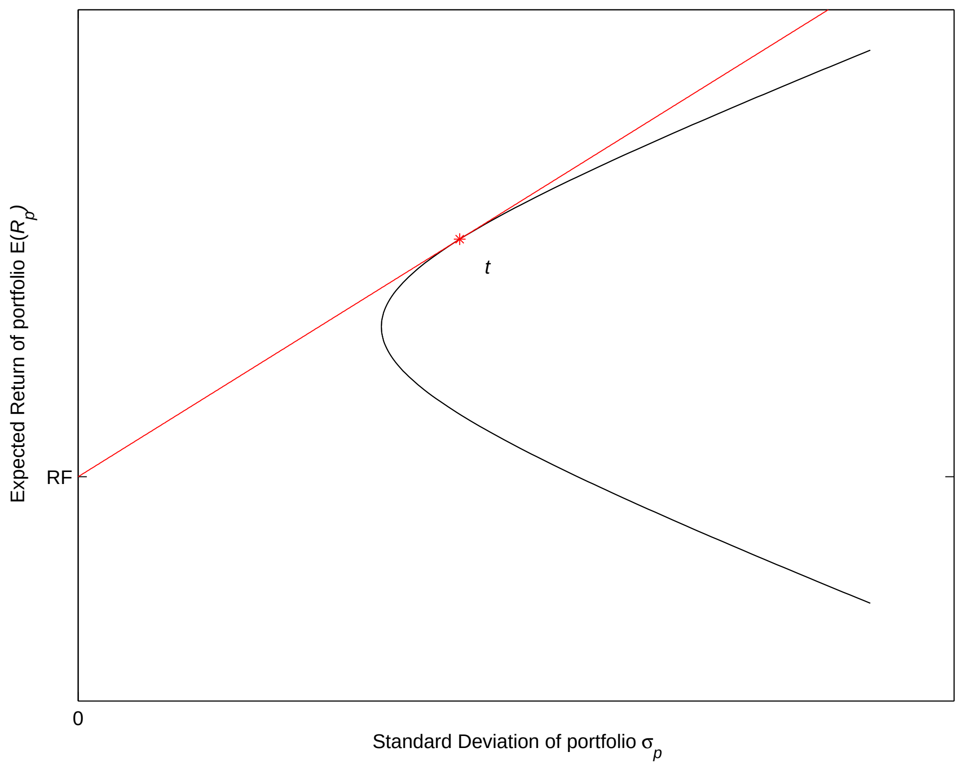

Figure 1: Efficient frontier for a market with a risk-free asset (straight line), and for a market without a risk-free asset (curve).

as the derivative of $\bar{R}_\phi$ with respect to σ_ϕ, evaluated at $\phi = 1$. Thus,

$$\left.\frac{\partial \bar{R}_\phi}{\partial \sigma_\phi}\right|_{\phi=1} = -\left.\frac{\frac{\partial \bar{R}_\phi}{\partial \phi}}{\frac{\partial \sigma_\phi}{\partial \phi}}\right|_{\phi=1} = \frac{\bar{R}_e - R_F}{\sigma_e}.$$

Solving for the above equation gives the result in (5).

The question we deal with in this paper is whether a given portfolio, called the benchmark, is efficient in the market $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$, where $\bar{R}_{+b} = \left(\bar{R}_1, \ldots, \bar{R}_N, \bar{R}_b\right)$ is an $(N+1) \times 1$ vector of mean returns of assets, where the $N+1$th asset is portfolio b, and Ω_{+b} is the variance-covariance matrix of the vector of returns R_{+b}. The question of the efficiency of the benchmark is clearly equivalent to asking whether it is true that

$$\bar{R}_n - R_F = \frac{\sigma_{bn}}{\sigma_b^2}\left(\bar{R}_b - R_F\right)$$

for $n \in \{1, \ldots, N\}$.

3 A Translation Family of Tests

Define $R_\gamma = R_F - \gamma$, where γ is an arbitrary real number. Efficiency of portfolio b can be re-stated to be the condition that

$$\bar{R}_n - R_\gamma = \frac{\sigma_{bn}}{\sigma_b^2}\left(\bar{R}_b - R_\gamma\right) + \left(1 - \frac{\sigma_{bn}}{\sigma_b^2}\right)\gamma \tag{6}$$

hold for $n \in \{1, \ldots, N\}$.

A regression-based empirical test of the efficiency of a benchmark portfolio, b, can be defined for every γ, thus defining a parametric family of efficiency tests.

3.1 Regression and Joint Null Hypothesis Test

In market $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$, let $\begin{bmatrix} R^T & R_F^T \end{bmatrix}$ be a sample of size T taken from the true distribution of asset returns, and let $R_\gamma^T = R_F^T - \gamma \mathbf{1}_T$, be the sample of R_F translated by γ. Fix γ and define the following projection of risky assets excess returns (in excess of R_γ) on the excess return of the benchmark portfolio:

$$\begin{aligned} & R_{n\tau} - R_{\gamma\tau} = \alpha_n(\gamma) + \beta_{\gamma n}(R_{b\tau} - R_{\gamma\tau}) + \epsilon_{n\tau} \\ & E[\epsilon_\tau] = \mathbf{0}, \quad E[\epsilon_\tau(R_{b\tau} - R_{F\tau})] = E[\epsilon_\tau(R_{b\tau} - R_{\gamma\tau})] = \mathbf{0} \\ & n \in \{1, \ldots, N\}, \quad \tau \in \{1, \ldots, T\}, \end{aligned} \tag{7}$$

where τ identifies the sample point or time, and $\epsilon_\tau = (\epsilon_{1\tau}, \ldots, \epsilon_{N\tau})$ is a vector of errors, assumed to be independent across τ.

If b is (mean-variance) efficient, equation (6) holds for all assets. The null hypothesis is thus a multivariate hypothesis on the value of $\alpha(\gamma) = (\alpha_1(\gamma), \ldots, \alpha_N(\gamma))$ and $\beta_\gamma = (\beta_{\gamma 1}, \ldots, \beta_{\gamma N})$. By construction, in the above projection, $\beta_{\gamma n} = \frac{\sigma_{bn}}{\sigma_b^2}$. Thus, the null hypothesis reduces to the hypothesis on the value of $\alpha(\gamma)$ given below:

$$\begin{aligned} H_0: \quad & \alpha(\gamma) = (1 - \beta_{\gamma n})\gamma \\ & \forall n \in \{1, \ldots, N\} \end{aligned} \tag{8}$$

Let $\hat{\alpha}(\gamma)$ and $\hat{\beta}_\gamma$ be the least-squares estimators of $\alpha(\gamma)$ and β_γ. Asymptotically, or under the additional assumption that ϵ_τ are normally distributed, $\hat{\alpha}(\gamma)$ satisfies

$$\hat{\alpha}(\gamma) \sim N\left(\alpha(\gamma), \frac{1}{T}\left[1 + \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}\right]\Sigma\right),$$

where

$$\begin{aligned} \hat{\mu}_{b\gamma} &= \frac{1}{T}\sum_{\tau=1}^{T}(R_{b\tau} - R_{\gamma\tau}), \text{ and} \\ \hat{\sigma}_b^2 &= \frac{1}{T}\sum_{\tau=1}^{T}(R_{b\tau} - R_{F\tau} - \hat{\mu}_b)^2 \end{aligned}$$

For expositional ease we will assume normality of the white noise error terms,

$$\epsilon_\tau \sim N(\mathbf{0}, \Sigma),$$

where Σ is the $N \times N$ matrix of covariances of $\epsilon_{n\tau}$ over n.

Given the least-squares regression estimates, $\hat{\alpha}(\gamma)$, there are several ways to test the joint null hypothesis, $\alpha_n(\gamma) = 0$ for all n. We follow Gibbons, Ross, and Shanken [1989], and propose the following test:

Definition 1 (γ-*GRS Test*)**.** We define the γ-GRS Test of the efficiency of a portfolio with the estimator

$$J(\gamma) = T\left[1 + \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}\right]^{-1} [\hat{\alpha}(\gamma)]' \, \Sigma^{-1} [\hat{\alpha}(\gamma)], \tag{9}$$

and the null hypothesis

$$H_0: \quad \alpha(\gamma) = (\mathbf{1} - \beta_\gamma)\,\gamma.$$

Under our projection assumptions, $J(\gamma)$ has a non-central chi-square distribution with N degrees of freedom and non-centrality parameter $\lambda(\gamma)$. That is,

$$\begin{aligned} &J(\gamma) \sim \chi_N^2(\lambda(\gamma)), \\ &\text{where } \lambda(\gamma) = T\left[1 + \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}\right]^{-1} \alpha(\gamma)' \, \Sigma^{-1} \alpha(\gamma). \end{aligned} \tag{10}$$

When Σ is unknown, it can be replaced with a consistent estimator $\hat{\Sigma}$. In that case, the finite-sample distribution of $J(\gamma)$ is F with N and $(T - N - 1)$ degrees of freedom (see Gibbons, Ross, and Shanken [1989]). In large samples $J(\gamma)$ will have a non-central χ^2 distribution, and the assumption of normality of ϵ_τ can be dropped.

We will work with the finite-sample (non-central) χ^2 distribution, but our analysis applies to the finite sample F distribution and the large-sample χ^2 distribution, *mutatis mutandis*. The statistical properties hold for all values of γ.

3.2 GRS Test; $\gamma = 0$

The γ-GRS test for $\gamma = 0$ deserves special attention. It is the original GRS test developed in Gibbons et al. [1989]. When $\gamma = 0$, the projection presented in (7), becomes a projection of assets' excess returns with respect to the risk-free return, on the excess return of the benchmark portfolio, b.

$$\begin{aligned} &R_{n\tau} - R_{F\tau} = \alpha_n(0) + \beta_{0n}(R_{b\tau} - R_{F\tau}) + \epsilon_{n\tau} \\ &\epsilon_\tau = (\epsilon_{1\tau}, \ldots, \epsilon_{N\tau}), \quad E[\epsilon_\tau] = \mathbf{0}, \quad E[\epsilon_\tau(R_{b\tau} - R_{F\tau})] = \mathbf{0}, \\ &n \in \{1, \ldots, N\}, \quad \tau \in \{1, \ldots, T\}. \end{aligned} \tag{11}$$

The test of efficiency is given by

$$J(0) = T\left[1+\frac{\hat{\mu}_{b0}^2}{\hat{\sigma}_b^2}\right]^{-1}\hat{\alpha}'(0)\,\Sigma^{-1}\hat{\alpha}(0)$$

$$H_0: \quad \alpha(0) = \mathbf{0}.$$

$J(0)$ has a non-central chi-square distribution with non-centrality parameter defined in (10). In this special case of $\gamma = 0$, the distribution of $J(0)$ under the null hypothesis is a chi-square with N degrees of freedom (the non-centrality parameter becomes zero).

The test estimator, $J(0)$ has a well known financial interpretation, that can be extended to $J(\gamma)$, for $\gamma \neq 0$. This interpretation, and an interpretation of the statistic's non-centrality parameter in terms of asset correlations, are the subject of the next section.

4 Interpretations of the γ-GRS Test

4.1 Sharpe Ratios

The GRS test has an insightful financial interpretation. The $J(0)$ statistic can be split in two parts. The term $T\left[1+\frac{\hat{\mu}_{b0}^2}{\hat{\sigma}_b^2}\right]^{-1}$ takes care of scaling, accounting for the size of the sample and the exact location of the benchmark portfolio. The term $\hat{\alpha}(0)'\,\Sigma^{-1}\hat{\alpha}(0)$ summarizes the estimated value of the intercept terms ($\alpha_n(0)$'s), accounting for the correlations between different securities. Gibbons, Ross, and Shanken [1989] show that

$$\hat{\alpha}(0)'\,\Sigma^{-1}\hat{\alpha}(0) = \left(\widehat{SR}^*\right)^2 - \left(\widehat{SR}^b\right)^2, \tag{12}$$

where $\widehat{SR}^*$ is the sample Sharpe ratio of the tangency portfolio in market $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$. $\widehat{SR}^b$ is the sample Sharpe ratio of the benchmark portfolio. $\widehat{SR}^*$ and $\widehat{SR}^b$ are defined to be

$$\left(\widehat{SR}^b\right)^2 = \frac{\hat{\mu}_{b0}^2}{\hat{\sigma}_b^2}, \text{ and}$$

$$\left(\widehat{SR}^*\right)^2 = \hat{\mu}_0'\hat{\Omega}_{+b}^{-1}\hat{\mu}_0,$$

where $\hat{\mu}_0$ and $\hat{\Omega}_{+b}$ are sample estimates of $\bar{R}_{+b}$ and Ω_{+b}. Thus,

$$\hat{\mu}_0 = \frac{1}{T}\sum_{\tau=1}^{T}\left(R_\tau - R_{F\tau}\mathbf{1}\right), \text{ and}$$
$$\hat{\Omega}_{+b} = \begin{bmatrix} \hat{\sigma}_b^2 & \hat{\sigma}_b^2\hat{\beta}_0' \\ \hat{\sigma}_b^2\hat{\beta}_0 & \sigma_b^2\hat{\beta}_0\hat{\beta}_0' + \Sigma \end{bmatrix}.$$

$J(\gamma)$, for $\gamma \neq 0$ can be analogously interpreted. Let $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$, be a market with the same set of risky assets as $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$, but where the risk-free return is given by $R_\gamma = R_F - \gamma$ instead of R_F. The theoretical and sample values of the Sharpe ratio of the tangency and the benchmark portfolio can be redefined for this market. The sample values are

$$\left(\widehat{SR}^b(\gamma)\right)^2 = \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}, \text{ and}$$
$$\left(\widehat{SR}^*(\gamma)\right)^2 = \hat{\mu}_\gamma'\hat{\Omega}_{+b\gamma}^{-1}\hat{\mu}_\gamma,$$

where

$$\hat{\mu}_\gamma = \frac{1}{T}\sum_{\tau=1}^{T}\left(R_\tau - R_{\gamma\tau}\right), \text{ and}$$
$$\hat{\Omega}_{+b\gamma} = \begin{bmatrix} \hat{\sigma}_b^2 & \hat{\sigma}_b^2\hat{\beta}_\gamma' \\ \hat{\sigma}_b^2\hat{\beta}_\gamma & \sigma_b^2\hat{\beta}_\gamma\hat{\beta}_\gamma' + \Sigma \end{bmatrix}.$$

The financial interpretation of $J(\gamma)$ is given by

$$\hat{\alpha}(\gamma)'\Sigma^{-1}\hat{\alpha}(\gamma) = \left(\widehat{SR}^*(\gamma)\right)^2 - \left(\widehat{SR}^b(\gamma)\right)^2.$$

In mean-standard deviation space, Sharpe ratios are the slopes of straight lines connecting the risk-free rate and the portfolios at hand. In the case of the GRS test, the question of efficiency of the benchmark portfolio can be interpreted as the question of whether the slope of the line passing through $(0, \hat{R}_F)$ and point $(\hat{\sigma}_b, \hat{\mu}_{b0})$, can be considered statistically equal to the slope of the sample efficient frontier. When $\gamma \neq 0$, the test statistic still compares the slopes of lines passing through the risk-free rate and either the tangency or the benchmark portfolio. However, the risk-free rate is now $\hat{R}_\gamma$, the tangency portfolio is efficient in market $\{\hat{\mu}_\gamma, \hat{\Omega}_{+b,\gamma}, \hat{R}_\gamma\}$, and the question is no longer whether these two slopes are equal. Instead, the question is whether the difference between the two slopes can be

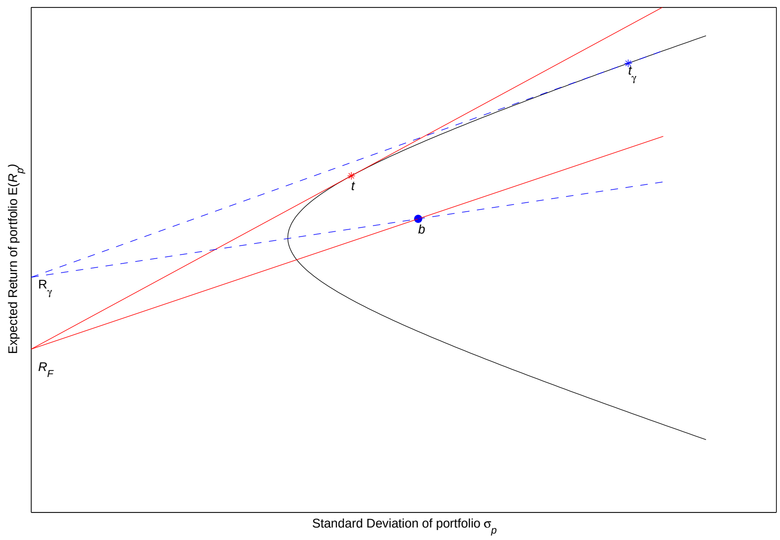

Figure 2: The solid lines are $\widehat{SR}^*$ and $\widehat{SR}^b$. The dashed lines are $\widehat{SR}^*(\gamma)$ and $\widehat{SR}^b(\gamma)$, for $\gamma \neq 0$.

deemed statistically equal to $\gamma^2 (\mathbf{1} - \beta_\gamma)' \Sigma^{-1} (\mathbf{1} - \beta_\gamma)$. The solid lines in Figure 2 are the Sharpe ratios $\widehat{SR}^*$ and $\widehat{SR}^b$. $\hat{\alpha}(0)' \Sigma^{-1} \hat{\alpha}(0)$ is the difference between the squared slopes of these two lines. The dashed lines are $\widehat{SR}^*(\gamma)$ and $\widehat{SR}^b(\gamma)$.

From the above discussion it is clear that $J(\gamma)$ is equal to $J(0)$ in market $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$, where the risk-free rate R_F is replaced with R_γ. In other words, in constructing $J(\gamma)$ we invent a fictitious market with a fictitious risk-free rate, R_γ, and a fictitious tangency portfolio - lying on the efficient frontier for $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$ - to which the benchmark portfolio is compared. If the null hypothesis were that $\alpha(\gamma) = \mathbf{0}$, this would be a test of efficiency of portfolio b in the market $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$. That is why the null hypothesis is appropriately modified to test efficiency of portfolio b in market $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$ using the tangency portfolio for market $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$.

It must also be noted that when we compare the slopes of the two dashed lines in Figure 2, we are in fact comparing the benchmark portfolio to a new reference portfolio t_γ (the tangency portfolio in market $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$) that is not necessarily efficient in market $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$. That is, the performance of the benchmark b will no longer be compared to the tangency portfolio t. If the new reference portfolio is not the best available portfolio, under the null that the benchmark portfolio is mean-variance optimal, the benchmark may sometimes outperform the (sub-optimal) reference portfolio. The adjusted critical region, through adjusted null hypothesis, will reflect this.

4.2 Portfolio Correlations

To better understand what statistical differences there are between $J(\gamma)$ for different values of γ, it is useful to interpret the population value $\alpha'(0)\,\Sigma^{-1}\alpha(0)$ in terms of the correlation between the benchmark and the tangency portfolios.[4]

Let $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$ be the market from which the sample $\begin{bmatrix} R^T & R_F^T \end{bmatrix}$ is taken. We will show that

$$\alpha_0'\Sigma^{-1}\alpha_0 = (SR^*)^2\left(1-\rho_{tb}^2\right), \tag{13}$$

where the absence of "hats" denotes population values, and ρ_{tb} is the correlation between the tangency portfolio t and the benchmark portfolio b, i.e.,

$$\rho_{tb} = \frac{\sigma_{bt}}{\sigma_t\sigma_b}.$$

To show (13), we must first note that the result in the previous subsection (see equation (12)) can be easily proven for population values. Thus,

$$\alpha_0'\Sigma^{-1}\alpha_0 = (SR^*)^2 - \left(SR^b\right)^2. \tag{14}$$

The tangency portfolio is efficient, and therefore,

$$\bar{R} - R_F\mathbf{1} = \frac{\Omega_{+b}w_t}{\sigma_t^2}(\bar{R}_t - R_F),$$

where $\frac{\Omega_{+b}w_t}{\sigma_t^2}$ is the vector of covariances of all assets in market $\{\bar{R}_{+b}, \Omega_{+b}, R_F\}$, with the tangency portfolio in that market. In particular, it is true that

$$\bar{R}_b - R_F = \frac{\sigma_{bt}}{\sigma_t^2}\left(\bar{R}_t - R_F\right). \tag{15}$$

Use equations (14) and (15) above, plus the definition of the Sharpe ratio of a portfolio

[4] Of course, the same interpretation holds for other values of γ. For illustration purposes it is more useful to limit attention to the interpretation of $\alpha'(0)\,\Sigma^{-1}\alpha(0)$, thus, the case of $\gamma = 0$.

to derive (13) as follows:

$$\begin{aligned}
\alpha_0'\Sigma^{-1}\alpha_0 &= (SR^*)^2 - \left(SR^b\right)^2 \\
&= (SR^*)^2 - \left[\frac{\frac{\sigma_{bt}}{\sigma_t^2}\left(\bar{R}_t - R_F\right)}{\sigma_b}\right]^2 \\
&= (SR^*)^2 - \left[\frac{\sigma_{bt}}{\sigma_t\sigma_b}\frac{\left(\bar{R}_t - R_F\right)}{\sigma_t}\right]^2 \\
&= (SR^*)^2 - \left(\frac{\sigma_{bt}}{\sigma_t\sigma_b}\right)^2 (SR^*)^2 \\
&= (SR^*)^2\left(1 - \rho_{bt}^2\right).
\end{aligned}$$

Under the null hypothesis that $\alpha(0) = \mathbf{0}$, the distribution of the GRS test estimator is central chi-square. However, equation (13) tells us that when the null hypothesis is not true, the non-centrality parameter of the test statistic will depend on the covariance between the (population) return of the benchmark portfolio and the tangency portfolio. This fact means that the power of the efficiency test depends on this correlation.

Specifically, the non-centrality parameter of the distribution of $J(0)$ equals

$$\lambda = T\left[1 + \frac{\left(\hat{\mu}_{b0} - R_F\right)^2}{\hat{\sigma}_b^2}\right]^{-1}\left[(SR^*)^2\left(1 - \rho_{tb}^2\right)\right]. \tag{16}$$

For a fixed critical region given by a point x_s, the area to the right under the relevant density is the power of the test. This area will normally be larger for a larger value of λ. From equation (16), *λ is larger for a smaller correlation between b and the tangency portfolio.*

Hence, to have a powerful test it is desirable to have a small correlation between the benchmark portfolio and the tangency portfolio. For a fixed benchmark portfolio, we will want to move the tangency portfolio to get better power. The use of a γ-GRS test is a way to take advantage of this possibility. By introducing $\gamma \neq 0$, we "move" the relevant tangency portfolio, and may improve power by doing this. The possibility of power improvement will be extensively dealt with in the next section.

5 Power of γ-GRS Tests

In this section we introduce new notation and definitions of power and size of the γ-GRS tests. These will be used in an example to illustrate the intuition behind the possibility of designing tests that are more powerful than the GRS test, and in our description of the power function at the end of the section.

Let $\lambda(\gamma)$ denote the non-centrality parameter of the distribution of $J(\gamma)$ when the null hypothesis is true ($\alpha_0 = \mathbf{0}$). If the null hypothesis is false, so that $\alpha(0) = c \neq \mathbf{0}$, we define

$$\tilde{\lambda}(\gamma, c) = T\left[1 + \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}\right]^{-1} \left[c + (\mathbf{1} - \beta_\gamma)\gamma\right]' \Sigma^{-1} \left[c + (\mathbf{1} - \beta_\gamma)\gamma\right],$$

which is the non-centrality parameter of the distribution of $J(\gamma)$ when $\alpha(0) = c$.

Let $F(y, \lambda)$ denote the probability that a variable with a $\chi_N^2(\lambda)$ distribution (non-central chi-square distribution with N degrees of freedom and non-centrality parameter λ) be smaller than y. Let $F^{-1}(s, \lambda)$ denote the sth quantile of a $\chi_N^2(\lambda)$ distribution.

Definition 2 (Critical Region)**.** The critical region of a statistical test is the set of values of the test statistic for which rejection of the null hypothesis is prescribed.

In the tests we consider here, the critical region is given by a cutoff point, x_s, such that the null hypothesis is rejected if $J(\gamma) > x_s$.

Definition 3 (Size)**.** The size of a test is the probability of rejecting the null hypothesis when it is true. Given a cutoff point x_s that defines a critical region $[x_s, \infty)$, the size of the γ-GRS test is given by

$$1 - F(x_s, \lambda(\gamma)). \tag{17}$$

It is common trade to fix the size of a test and find the critical region that will deliver this size. That is, given a size, $1 - s$, a cutoff point

$$x_s(\gamma) = F^{-1}(s, \lambda(\gamma)) \tag{18}$$

can be found to define the critical region for the test. In equations (17) and (18) $\lambda(\gamma)$ is given by

$$\lambda(\gamma) = T\left[1 + \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}\right]^{-1} \gamma^2 (\mathbf{1} - \beta_\gamma)' \Sigma^{-1} (\mathbf{1} - \beta_\gamma),$$

which is the value of the non-centrality parameter for the distribution of $J(\gamma)$ under the null hypothesis. The power of a test relates to the distribution of the test statistic under the alternative hypothesis. Since the alternative hypothesis is typically an interval, not a point, power is given by a function over this interval. We will return to this in a later subsection.

Definition 4 (Power)**.** The power of a test is the probability of rejecting the null hypothesis when it is false. Given a critical region $[x_s, \infty)$, and given the true value of the Jensen α's, $\alpha(0) > 0$, in market $\{\bar{R}_{+b}, \Omega_{+b}, R_\gamma\}$, the power of the γ-GRS test is given by

$$1 - F\left(x_s, \tilde{\lambda}(\gamma, \alpha(\gamma))\right). \tag{19}$$

The following example illustrates the way in which the power and the size of the γ-GRS test relate to the value of γ.

5.1 Example: Power Changes with Value of γ

The example we give in this subsection shows a dramatic (unfavorable) change in the power of an efficiency test, as a product of using a specific $\gamma \neq 0$, instead of the standard $\gamma = 0$ (GRS test). In this exercise, no attempt is made to maximize power. The value of γ is chosen in the "clumsiest" possible way, to show how badly things can go if γ is poorly picked.

Consider a situation where the benchmark portfolio b is *not* efficient, but lies on the efficient frontier for the market with risky assets only, $\{\bar{R}_{+b}, \Omega_{+b}\}$. In this case, a γ^b can be found, such that the combinations of risky assets and a fictitious risk-free asset with return $R_{\gamma^b} = R_F - \gamma^b$, lie on the line tangent to the efficient frontier for market $\{\bar{R}_{+b}, \Omega_{+b}\}$ at b. Figure 3 illustrates this situation.

The null hypothesis ($H_0 : \alpha(0) = \mathbf{0}$) can be tested using $J(\gamma)$ for any γ. We use two tests from this family and compare their performance in terms of power. The tests we choose are $J(0)$, and $J\left(\gamma^b\right)$. We also pick the tests to have the same size $1 - s$ (in the figures we set $1 - s = 0.05$). *If the null had been true*, the distributions of $J(R_F)$ and $J\left(\gamma^b\right)$ would be given by the non-centrality parameters $\lambda(0)$ (always equal to zero), and $\lambda\left(\gamma^b\right)$ respectively.

	$\boldsymbol{\gamma = 0}$	$\boldsymbol{\gamma = \gamma^b}$
$\boldsymbol{\alpha(\gamma)}$	$\mathbf{0}$	$\left(\mathbf{1} - \beta_{\gamma^b}\right)\gamma^b$
$\boldsymbol{\lambda(\gamma)}$	0	$T\left[1 + \frac{\hat{\mu}_{b\gamma}^2}{\hat{\sigma}_b^2}\right]^{-1}\left(\gamma^b\right)^2(\mathbf{1} - \beta_\gamma)'\,\Sigma^{-1}(\mathbf{1} - \beta_\gamma)$

However, *the null is not true.* In this case, we know that $\alpha_0 = -\gamma^b\left(\mathbf{1} - \beta_{\gamma^b}\right)$ (since $\alpha\left(\gamma^b\right) = 0$ because b is an efficient portfolio in the fictitious market where R_F is replaced with R_{γ^b}, $\{\bar{R}_{+b}, \Omega_{+b}, R_{\gamma^b}\}$). As a result, under the *true* alternative hypothesis that $\alpha_0 = -\gamma^b\left(\mathbf{1} - \beta_{\gamma^b}\right)$, the parameter $c(\gamma)$ (alternative hypothesis value of $\alpha(\gamma)$) and the non-centrality parameter are given by

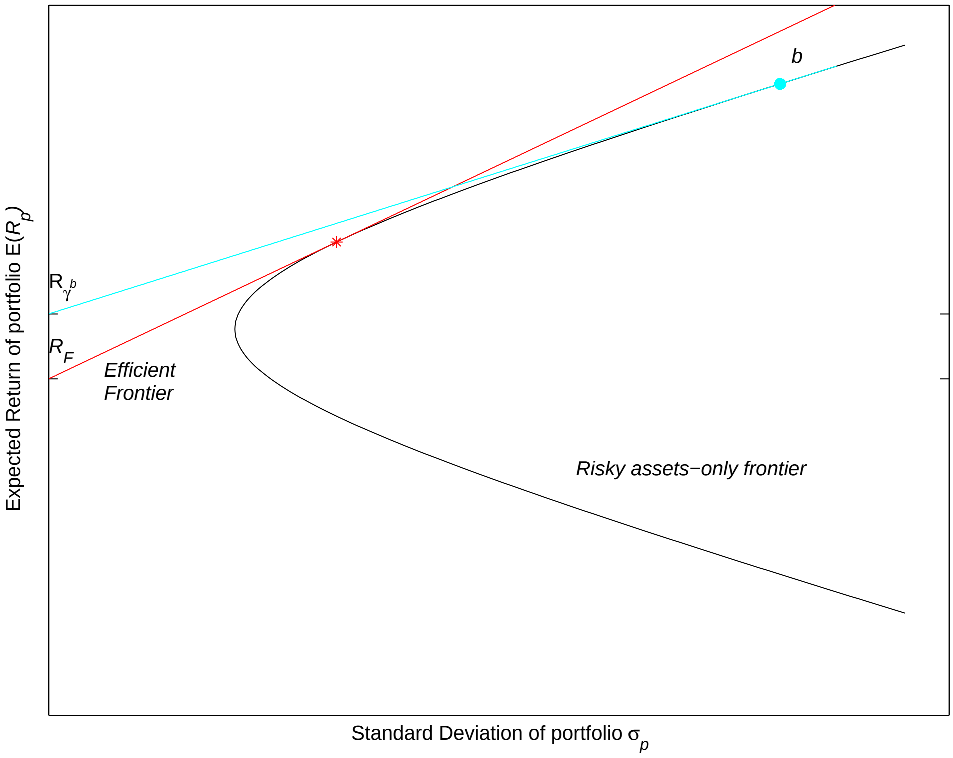

Figure 3: The benchmark portfolio b is not efficient, but lies on the efficient frontier for market $\{\bar{R}_{+b}, \Omega_{+b}\}$. We can draw a tangent that has intercept R_{γ^b}.

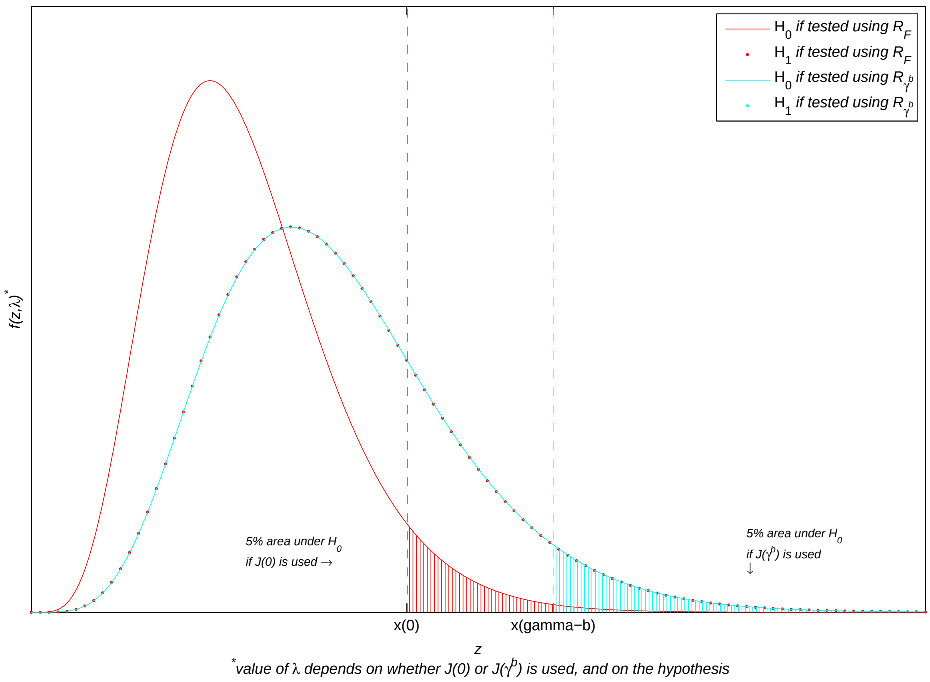

Figure 4: At the significance level 0.95, critical regions are displayed for $J\left(\gamma^b\right)$ and $J\left(0\right)$.

	$\boldsymbol{\gamma = 0}$	$\boldsymbol{\gamma = \gamma^b}$
$\boldsymbol{c\left(\gamma\right)}$	$-\gamma^b\left(\mathbf{1} - \beta_{\gamma^b}\right)$	$\mathbf{0}$
$\boldsymbol{\tilde{\lambda}\left(\gamma, c\right)}$	$T\left[1 + \frac{\hat{\mu}_b^2}{\hat{\sigma}_b^2}\right]^{-1}\left(-\gamma^b\right)^2\left(\mathbf{1} - \beta_{\gamma^b}\right)'\Sigma^{-1}\left(\mathbf{1} - \beta_{\gamma^b}\right)$	0

The combination of the fixed size $1 - s$ and the distribution under the false null hypothesis may produce the situation depicted in Figure 4.[5]

The *cdf* of $J\left(0\right)$ under the (false) H_0 coincides with the *cdf* of $J\left(\gamma^b\right)$ for the (true) H_1. Critical regions for both tests are determined using the distribution under the (false) null hypothesis. These regions, for size 0.05 are depicted in figure 4. The cutoff point for $J\left(0\right)$ lies to the left of the cutoff point for $J\left(\gamma^b\right)$, defining a probability of rejecting the null hypothesis, given that it is false, which is depicted in Figure 5. The *cdf* of $J\left(\gamma^b\right)$ under the (false) H_0 is almost identical to that of $J\left(R_F\right)$ under H_1.[6] In Figure 5 we can immediately see that *the test using* $J\left(0\right)$ *will be more powerful.*

[5] We say "may" because the non-central χ^2 is not linear in the non-centrality parameter. The exact values of $\lambda\left(R_F\right)$ and $\lambda\left(\gamma^b\right)$, the degrees of freedom, and the size will jointly determine the ordering of cutoff points.

[6] To generate these plots we used CRSP data compiled by Fama and French [2004]. The specific series are monthly returns to 12 industry portfolios, returns from one-month Treasury Bills, and a portfolio of

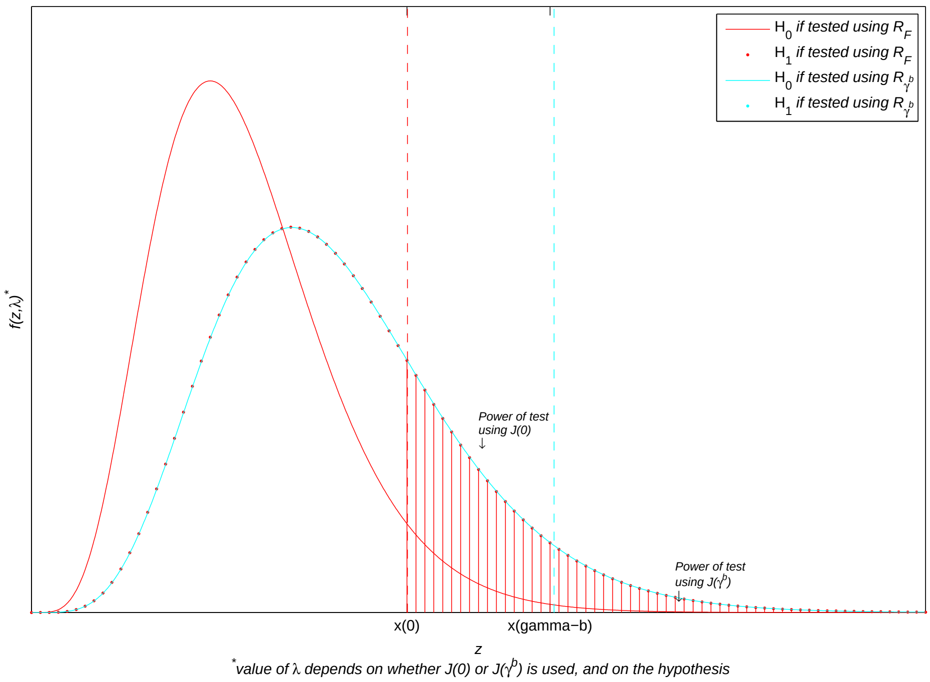

Figure 5: Power of the tests $J\left(\gamma^b\right)$ and $J\left(0\right)$ with size 0.05.

We will now attempt to use the potential depicted in this example to our advantage. The objective is to design a test that is more powerful than the GRS test.

5.2 Power Function and Power Improvement

Fix a critical region given by the cutoff point x_s. For each γ we can define the power function given the critical region to be

$$p\left(c; x_s, \gamma\right) = P\left(J\left(\gamma\right) > x_s | \alpha\left(0\right) = c\right) = 1 - F\left(x_s, \tilde{\lambda}\left(\gamma, c\right)\right), \tag{20}$$

where c is an arbitrary non-zero vector. That is, c is a value of $\alpha\left(0\right)$ that does not satisfy the null hypothesis. The power of a test depends on the exact *true* value of the relevant

indices. The benchmark b is chosen to fit the example. Even though it seems that the *cdf* of $J\left(\gamma^b\right)$ under H_0 is identical to that of $J\left(0\right)$ under H_1, they are slightly different. In particular: $\lambda\left(\gamma^b\right) = 4.95$, while

$$\tilde{\lambda}\left(0, -\gamma^b\left(\mathbf{1} - \beta\right)\right) = 4.93.$$

This difference is introduced by the scaling factor $T\left[1 + \frac{\left(\hat{\mu}_{b\gamma^b}\right)^2}{\hat{\sigma}_b^2}\right]^{-1}$.

parameter ($\alpha(0)$, in this case). Since this value is unknown, power functions are defined for all possible values of the relevant parameter that correspond to the alternative hypothesis.

Equation (20) gives the power for a fixed cutoff point, and different values of γ and c. Notice, though, that the size of the test for a fixed cutoff point, is also changing with γ. It is more desirable to fix the size of the test, allow the cutoff point to adjust appropriately (as a function of γ), and define power as a function of size, γ, and the value c.

$$\pi(c; s, \gamma) = 1 - F\left(x_s(\gamma), \tilde{\lambda}(\gamma, c)\right), \tag{21}$$

where $x_s(\gamma) = F^{-1}(s, \lambda(\gamma))$. Hence, the power function is affected by the distribution both under the null and under the alternative hypotheses.

The naive approach to the problem of finding an optimal test in the γ-family of tests, is to attempt to maximize the power function stated above. This function is differentiable and locally concave, so first order conditions can be found (and second order conditions locally checked), and are stated below.

FOC:

$$\begin{aligned} 0 = & \frac{\partial F\left[F^{-1}(s, \lambda(\gamma)), \tilde{\lambda}(\gamma, c)\right]}{\partial \tilde{\lambda}(\gamma, c)} \frac{\partial \tilde{\lambda}(\gamma, c)}{\partial \gamma} \\ & + \frac{\partial F\left[F^{-1}(s, \lambda(\gamma)), \tilde{\lambda}(\gamma, c)\right]}{\partial F^{-1}(s, \lambda(\gamma))} \frac{\partial F^{-1}(s, \lambda(\gamma))}{\partial \lambda(\gamma)} \frac{\partial \lambda(\gamma)}{\partial \gamma}. \end{aligned}$$

Two points become immediately apparent from setting up the maximization problem:

1. The dependence on the alternative hypothesis, c, does not disappear at the maximum. That is, there is no γ that will maximize power for all possible points violating the null hypothesis.[7]

2. The first order condition cannot be solved analytically, even for a fixed value of c. Numerical methods must be use to find maxima.

3. The fact that there is no test that maximizes power for all values of c implies that typically the GRS test will not maximize power. Moreover, the GRS needn't maximize power at the *in-sample* value of portfolio correlations ($\hat{\sigma}_{bt}$).

[7]Notice that the notion of uniformly most powerful test in this case is different from the standard use of the term. This is so because the possibility of affecting power through γ relies on an assumption about the correlation between the benchmark portfolio and t_γ, which is an assumption about c, the true distribution of asset returns. In other words, it makes no sense to expect there to be a uniformly most powerful test in this situation.

	R_1	R_2	R_3	R_4	R_5	R_6	R_7	R_8	R_9	R_{10}	R_{11}	R_{12}
$\widehat{\beta}_n$	0.77	1.17	1.19	0.85	0.98	1.29	0.66	0.81	0.96	0.86	1.13	1.14
$\widehat{\sigma}_\beta$	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
$\widehat{\alpha}_{0n}$	0.19	0.12	-0.07	0.18	0.13	0.01	0.13	0.05	0.08	0.28	0.04	-0.18
$\widehat{\sigma}_{\alpha_0}$	(0.08)	(0.12)	(0.07)	(0.12)	(0.09)	(0.11)	(0.1)	(0.12)	(0.09)	(0.11)	(0.09)	(0.09)

Table 1: Regression (LSE) values of the parameters α_0 and β

In the absence of a γ that maximizes power for all return distributions in the alternative hypothesis, there are several options of differing validity. The approach we take in the remainder of the paper is to use a sample estimate of the true value of c and numerically determine the power maximizing γ for this value. There are several advantages of this approach, starting with its simplicity, and its asymptotic validity (plus the fact that it does not affect asymptotic distributions of the relevant statistics). Finally, for the purpose of our examples, this approach serves to illustrate that under the (implicit in all statistics) assumption that the estimated values are close to true, the GRS can have very poor power. The main disadvantages of this approach is that it uses sample information "twice", and it limits all results to one point, with no consideration of robustness. In the following section we use this approach to illustrate the potential power gains from γ-GRS tests.

6 Numerical Examples

The purpose of this section is to demonstrate that power improvements can indeed be achieved by using $\gamma \neq 0$. We use time series of CRSP monthly returns on 12 value-weighted industrial portfolios, taken from Fama and French [2004]. The benchmark is also taken from the Fama and French database, and is a weighted combination of the returns on NYSE, AMEX, and NASDAQ securities. We take the returns on 1-month Treasury Bills to be the risk-free rate. All series run from July 1926 to December 2003 and are given in percentage points.

We fix three levels of significance s and find the optimal test by replacing the unknown parameters with estimates. The first step of this procedure is to run the regressions specified in (11). Sample values of $\hat{\beta}$, $\hat{\Sigma}$, and $\hat{\alpha}(0)$ are taken from these regressions and inserted in the power function, which is subsequently used to find the optimal value of γ. The least squares regression coefficients and corresponding standard deviations for each of the industrial portfolios are reported in Table 1.

Table 2 reports the power maximizing γ for three levels of significance: 0.9, 0.95, and

s	R_F	$\bar{R}_\gamma$	$100 \times \frac{\pi(\hat{\alpha}(0);s,\bar{\gamma})-\pi(\hat{\alpha}(0);s,0)}{\pi(\hat{\alpha}(0);s,0)}$
0.99	0.31	0.05	3.75
0.95	0.31	0.02	1.3
0.90	0.31	0.005	0.64

Table 2: Overall maximal power values for sample 1926-2002. $\hat{R}_\gamma$ is the fictitious risk-free rate that corresponds to the power maximizing γ. The gain in power from using the power maximizing γ is also reported.

0.99 when estimation is made over the entire sample. The average risk-free rate over the sample period is reported for comparison. The percentage increase in power of the test, when the optimal γ (denoted $\bar{\gamma}$) is used instead of the risk-free rate, is given in the last column. The reported risk-free rate is the average monthly percentage return on T-bonds. Therefore, if we think of R_γ as a fictitious risk-free rate, the reported $\bar{R}_\gamma$ will be monthly percentage returns.

Figure 6 illustrates estimated power as a function of γ for $s = 0.99$, for the entire sample. The vertical line marks the location of R_F (since the return on T bills varies over the sample, we take R_F to be the average return on T bills) and crosses the power curve at $\pi(\hat{\alpha}_0, R_F, s)$. The maximal value of $\pi(\hat{\alpha}(0); s, \gamma)$ is achieved for $\bar{R}_\gamma = R_F - \bar{\gamma}$ smaller than R_F ($\bar{R}_\gamma = 0.05$).

Figure 7 shows the values of $\bar{R}_\gamma$ and R_F for 5-year periods starting June of each year comprised between 1926 and 2002. The test has a significance $s = 0.99$, and hence its size is $1 - s = 0.01$. Figure 8 shows the percentage power gains achieved by using $R_{\bar{\gamma}}$ instead of R_F over the same group of 5-year periods.

The percentage power gain over the considered five-year periods peaks for the period starting in 1936, when it is 22.4%. This means that the probability of falsely accepting the mean-variance efficiency of the benchmark is reduced by as much as 18%.

7 An Open Question and Extensions

7.1 An Open Question: The Gibbons Test

The Gibbons test of efficiency is an alternative to the GRS test, which has in occasions delivered rejections of the null hypothesis when the GRS does not reject. This test assumes there is no risk-free rate, and is expected to be lenient to deliver results that do not reject the null hypothesis. It is meant for markets where there is no risk-free asset, or where this asset is unknown. When there is no risk-free rate, a result like that of equation (6) can be

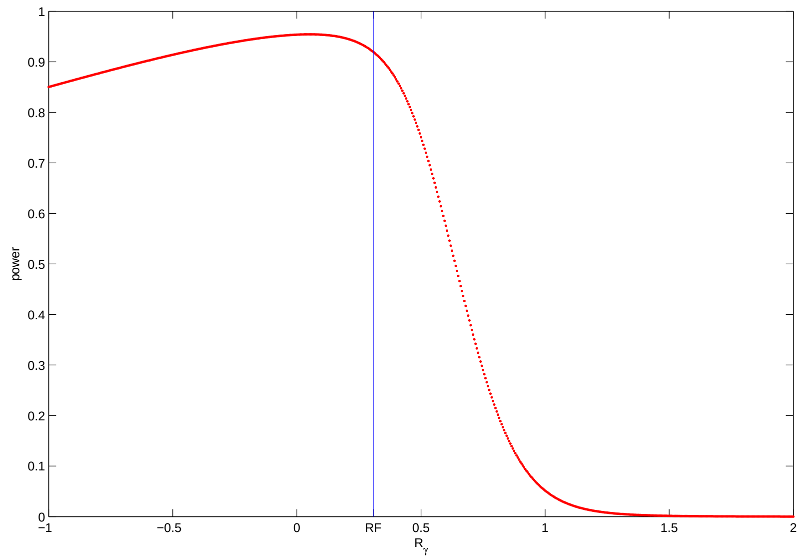

Figure 6: Monthly CRSP data on 12-industry portfolios, 1926-2002. Given $c = \hat{\alpha}_0$, and $s = 0.99$, we can find power as a function of R_γ.

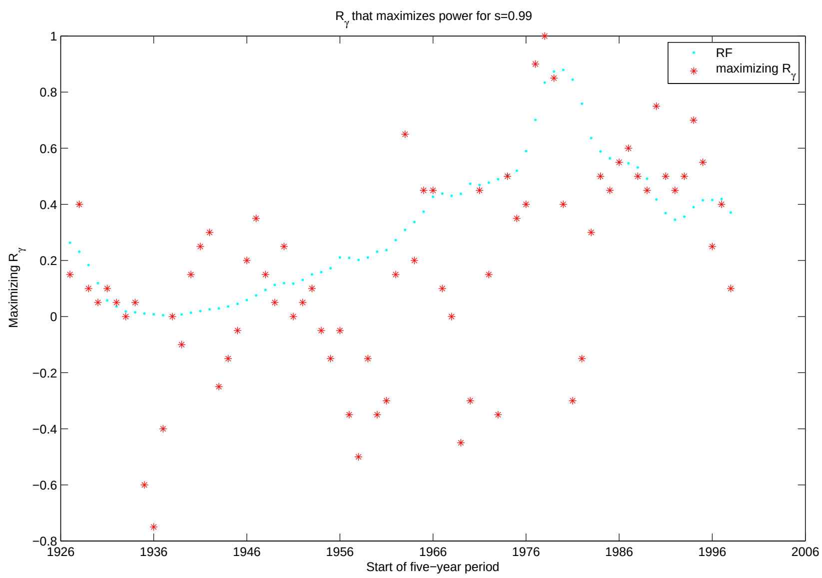

Figure 7: Yearly 5-year periods between 1926 and 2003. The average risk-free rate and power maximizing R_γ for each period are plotted at the beginning of the period.

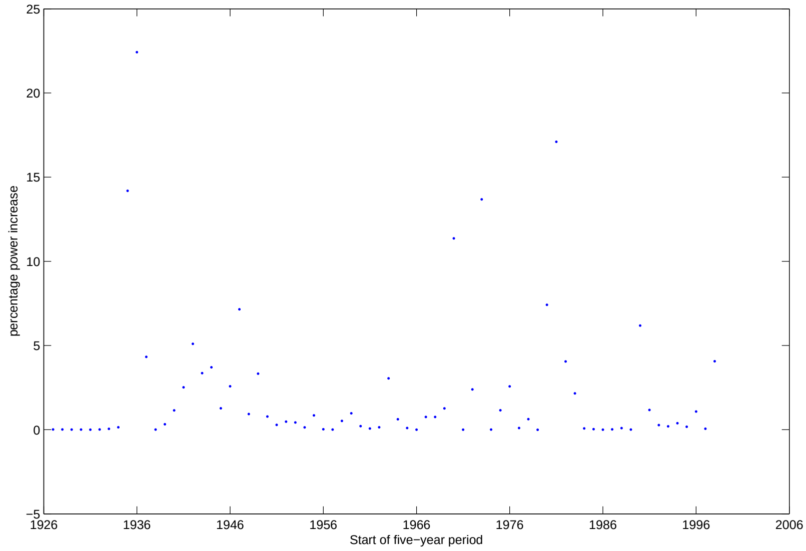

Figure 8: Percentage increase in power achieved using maximizing R_γ instead of R_F for yearly five-year periods between 1926 and 2002. The percentage increase corresponding to each period is plotted at the beginning of the period.

stated.

$$\bar{R}_n - \bar{R}_{zb} = \frac{\sigma_{bn}}{\sigma_b^2}\left(\bar{R}_b - \bar{R}_{zb}\right), \tag{22}$$

where zb denotes b's zero-covariance portfolio, which is a portfolio for which $\sigma_{b,zb} = 0$. This zero-covariance portfolio is unobservable, and $\bar{R}_{zb} = \delta$ becomes a parameter of the model. Gibbons proposes the following econometric model to estimate δ:

$$\begin{gathered} R_{n\tau} - \delta = \alpha_n(\delta) + \beta\left(R_{b\tau} - \delta\right) + \epsilon_{n\tau} \qquad (23) \\ \epsilon_\tau = \left(\epsilon_{1\tau}, \ldots, \epsilon_{N\tau}\right); \quad \epsilon_\tau \sim N\left(\mathbf{0}, \Sigma\right) \\ n \in \{1, \ldots, N\}, \quad \tau \in \{1, \ldots, T\} \end{gathered}$$

The test of mean-variance efficiency of the benchmark portfolio then becomes a test of *whether there exists a* δ *such that*

$$\alpha(\delta) = \left(\alpha_1(\delta), \ldots, \alpha_N(\delta)\right) = \mathbf{0}. \tag{24}$$

Parameter δ can be estimated using maximum likelihood. Let δ^* denote the maximum likelihood estimate. A statistic similar to $J(\delta)$ is then computed. Denote it $J^*(\delta)$. One difference with the regression for the γ-GRS with $\gamma = \delta$ is that the fictitious risk free rate is fixed, while R_δ varies across sample points. This affects the estimated variance of the assets and portfolios. If this effect is small (for example, if the risk-free rate R_F does not vary much in time), the following discussion sheds light on the differences between the GRS and the Gibbons tests.

The main difference between the Gibbons test with intercept δ^* and $J(\delta^*)$ is that in the former the acceptance region for size $1 - s$ is determined using a *central* χ^2_N distribution instead of $\chi^2_N(\lambda(\delta^*))$.[8] As a result, the Gibbons test uses the wrong critical region if indeed a risk-free rate exists. That is, its size is wrong. At the same time, since it uses a central χ^2_N distribution instead of the $\chi^2_N(\lambda(\delta^*))$, given δ^*, the power of the Gibbons test is larger than that of $J(\delta^*)$. The question remains of whether the criterion of selection of δ^* naturally leads to a correlation between the benchmark and the tangency portfolio in the market with risk-free rate δ^* that improves power. This is a relevant question we leave for future exploration.

7.2 Extensions

The main extension relates to the applicability of the power maximization procedure we propose and illustrate in the numerical examples. Confidence intervals for the power maximizing γ must be found and can be found through bootstrapping.

A second extension is to attempt maximization of power over subsets of the alternative hypothesis or points that are defined more abstractly than the one we use now. One such point is the null hypothesis itself. That is, one can ask the question of what γ will maximize power of the efficiency test when the truth approaches the null hypothesis. This is the *locally most powerful test* approach. The option we chose in this paper is to have a point-wise most powerful test, which is particularly useful for illustration of the power properties of the GRS test. It may be more useful in practice to have more general results regarding the direction of change in γ leading to power improvements.

References

Eugene Fama and Kenneth French. Data library, 2004. URL http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

[8] This is an abuse of notation. When we presented the family of tests $J(\gamma)$, γ referred to the difference between the fictitious and the true risk-free rate in the economy. Here, γ is replaced with δ^*, but δ^* is the fictitious risk-free rate, not the difference between the dummy and the real rate.

Michael Gibbons. Multivariate tests of financial models: A new approach. *Journal of Financial Economics*, 10:3–27, 1982.

Michael Gibbons, Stephen Ross, and Jay Shanken. A test of the efficiency of a given portfolio. *Econometrica*, 57(5):1121–1152, 1989.