Yes, the CAPM is Testable

Abstract

It is well-known that cross-sectional tests of the CAPM are problematic. The market indexes used in empirical tests are likely to be inefficient ex ante, which could lead to spurious results even in the absence of sampling errors. This problem has led many to express serious doubt on the testability of the CAPM. In this paper I show that the CAPM is indeed testable. This paper builds on the seminal paper by Kandel and Stambaugh (1995) and proposes a two-step procedure for testing the CAPM. The first step uses a simple combination of the coefficients of determination from both Ordinary Least Squares and Generalised Least Squares estimations. This step tests whether the index used in the empirical test is efficient and whether there are no omitted factors. The second step tests the hypothesis that the efficient index is the market portfolio. The two-step approach enables testing the CAPM regardless of whether the true expected return generating process is a CAPM or a non-CAPM.

1. Introduction

One of the central results of the Sharpe (1964), Lintner (1965) and Black (1972) (henceforth SLB) Capital Asset Pricing Model (CAPM) is that the relationship between beta and expected returns is linear, exact, and has a slope equal to the expectation of the (efficient) market portfolio excess return. Despite its appeal, this prediction has found little support in more than three decades of empirical studies. This is probably not surprising. Empirical tests usually employ observed variables, while none of the variables involved in empirical tests of the CAPM are directly observable. Cross-sectional tests of the CAPM involve replacing true variables with estimates and proxies. Expected returns are proxied by average realized returns, betas are estimated from a sample of time series returns, and the market portfolio is proxied by an index. Averaging realized returns and estimating betas introduce estimation error and, hence, affect the power of empirical tests. However, they do not change the fundamental behaviour of the SLB model. On the other hand, employing indexes other than the true market portfolio appears to fundamentally change the behaviour of the SLB model (Roll, 1977; Roll and Ross, 1994; and Kandel and Stambaugh, 1995).

Kandel and Stambaugh (1995), henceforth KS, found that Ordinary Least Squares (OLS) estimates of beta-return relation bear no relation to the proximity of an inefficient index to the efficient frontier. More importantly, KS show that while OLS estimates are of little use, Generalised Least Squares (GLS) estimates are functions of known quantities that are exactly related to the position of the index within the mean-variance frontier. These can be used to assess how close the used index is to the efficient frontier. But while the GLS quantities derived in KS are useful for testing the efficiency of a particular index, they are not useful for testing the CAPM. The first aim of this paper is to extend the results of KS and show that, if the CAPM is the true generating process, these GLS quantities can be used to test it.

KS investigated the effect of using an inefficient proxy on the CAPM under the assumption that some version of the CAPM holds. This assumption may not be correct, as expected returns may well be described by, say, some unknown multifactor model. In such a case, the KS results no longer hold because we now have an additional source of misspecification, namely the omission of some relevant risk factor(s). The second aim of this paper is to show that there is a simple way to test the CAPM even when the assumption that the CAPM is the true generating process is incorrect. I achieve this by using the coefficient of determination from both OLS and GLS methods, which can help distinguish cases of inefficient indexes from cases of omitted unobserved variables.

This paper is closely related to Grauer and Janmaat (2009) and Murtazashvili and Vozlyublennaia (2012). These authors provide useful simulation results and briefly review the literature related to what Grauer and Janmaat (2009) call the first stream of literature, which is mainly concerned with the power of empirical tests of the CAPM. However, while the present paper has empirical implications, it falls within what Grauer and Janmaat (2009) call the second stream, which focuses on population parameters and the fundamental economic behaviour of the CAPM.

The rest of the paper begins with a summary of the full procedure proposed in this paper in Section 2. In Section 3 the true cross-sectional relation between expected returns and beta computed from any index is outlined. Section 4 removes potential ambiguity by giving a clear idea of what is meant by a 'correctly specified' or 'true' model. In Section 5 I show how we can distinguish the true market index from any other efficient index, and provide a direct test of the CAPM under the assumption that the CAPM is the true generating process. Section 6 deals with potential non-CAPM models, and proposes a simple procedure to simultaneously assess the efficiency of the index and the presence of omitted factors. Section 7 provides a simulation experiment to demonstrate the behaviour of the OLS and GLS statistics. The final section concludes.

2. Summary of the Proposed Test

In this paper I will employ the following notation, which is similar to the one used in Roll and Ross (1994), henceforth RR,[1]

$\mathbf{R}$ is the expected returns vector for N individual assets,

$\mathbf{V}$ is the $N \times N$ covariance matrix of returns,

$\mathbf{1}$ is the unit vector,

Bold lower case letters ($\mathbf{i}$, $\mathbf{m}$, $\mathbf{q}$ and $\mathbf{p}$) are portfolio weights vectors,

$\boldsymbol{\beta}_i = (\beta_{1i}\ \beta_{2i}\ \cdots \beta_{Ni})' = \mathbf{Vi}/\sigma_i^2$,

$r_i = \mathbf{i}'\mathbf{R}$ is the expected return of portfolio $\mathbf{i}$,

$\sigma_i^2 = \mathbf{i}'\mathbf{Vi}$ is the variance of portfolio $\mathbf{i}$.

The testing procedure I propose is simple. It is based on the idea that both OLS and GLS R^2 will only attain the maximum value of one if there were no omitted factors and the index used to compute beta is efficient.

A general true process generating expected returns can be written as

$$\mathbf{R} = r_{zm}\mathbf{1} + \gamma_p \boldsymbol{\beta}_p + \mathbf{e}_p + \mathbf{d}$$

where r_{zm} is the true market portfolio's zero beta rate, γ_p is a parameter, and the vector $\boldsymbol{\beta}_p$ is computed against a possibly inefficient portfolio $\mathbf{p}$. Even

[1] Like RR and KS, I assume that expected returns and the covariance matrix of these returns are known.

without measurement errors, cross-sectional tests of the CAPM have to contend with two unobserved error terms. Both terms are generally correlated with the regressor ($\boldsymbol{\beta}_p$) and hence lead to systematic biases. The first error term, $\mathbf{e}_p$, is due to possible inefficiency of the index used to compute $\boldsymbol{\beta}_p$. The second term, $\mathbf{d}$, is due to the possibility that the CAPM is not the true generating process.

If expected returns are explained by an efficient portfolio beta, then $\mathbf{d} = \mathbf{0}$. We can only ensure that $\mathbf{d} = \mathbf{0}$ if both OLS and GLS R^2 attain their maximum value. The first step, therefore, in testing the CAPM would involve testing the null

$$H_{01} : R^2_{gls} = R^2_{ols} = 1.$$

If this hypothesis is accepted, this would confirm that some form of beta model holds and that the index used in computing beta is efficient. The next step consists of testing the hypothesis that the sum of GLS-estimated slope, $\gamma_p(gls)$, and intercept, $r_{zp}(gls)$, are equal to the expected index return, r_p. In other words, the second step tests the following null

$$H_{02} : r_{zp}(gls) + \gamma_p(gls) - r_p = 0$$

Failing to reject this second hypothesis would confirm that the index is indeed the market portfolio and that, consequently, the CAPM holds.

3. Inefficient Indexes and the True Relation between Expected Returns and Beta

In this section I offer a simple representation of the relation between beta computed on a inefficient index and expected returns. If the SLB model holds, there exists an efficient market portfolio, $\mathbf{m}$, for which expected returns have the following exact linear relation

$$\mathbf{R} = r_{zm}\mathbf{1} + \gamma_m \boldsymbol{\beta}_m \tag{1}$$

where r_{zm} is portfolio $\mathbf{m}$'s zero beta rate, and $\gamma_m = r_m - r_{zm}$ is the risk premium. The vector $\boldsymbol{\beta}_m = (\beta_{1m}\ \beta_{2m}\ \cdots \beta_{Nm})'$ is computed against the efficient market portfolio $\mathbf{m}$. If we were to observe $\boldsymbol{\beta}_m$, model (1) can be used to find the exact zero beta rate and the risk premium. However, given $\mathbf{V}$, observing $\boldsymbol{\beta}_m$ is dependent upon observing the efficient portfolio $\mathbf{m}$. In practice, though, it is almost impossible to observe $\mathbf{m}$. Suppose that instead we use an observable, but possibly inefficient, index $\mathbf{p}$. Two questions arise: (i) what would the true relation between beta computed on the observed index and expected returns be; and (ii) can we still test the CAPM given that we are using possibly the wrong index? The answer to the last question is yes. This is achieved by first finding out whether or not we are using an efficient index and then proceeding to assess whether the efficient index is the market portfolio. This is discussed in subsequent sections. Before that, I will attempt to answer the first question.

We can write the possibly inefficient portfolio $\mathbf{p}$ as a deviation from the efficient market portfolio, that is

$$\mathbf{m} = \mathbf{p} + \mathbf{D}_p \tag{2}$$

In this identity, the vector $\mathbf{D}_p$ is simply the difference between $\mathbf{m}$ and $\mathbf{p}$. The only constraint is that $\mathbf{D}_p$ is a zero investment portfolio ($\mathbf{D}_p'\mathbf{1} = 0$).

Replacing in (1) gives

$$\begin{aligned}\mathbf{R} &= r_{zm}\mathbf{1} + \frac{\gamma_m}{\sigma_m^2}\mathbf{Vm} \\ &= r_{zm}\mathbf{1} + \frac{\gamma_m}{\sigma_m^2}\mathbf{Vp} + \frac{\gamma_m}{\sigma_m^2}\mathbf{VD}_p \\ &= r_{zm}\mathbf{1} + \gamma_p\boldsymbol{\beta}_p + \mathbf{e}_p \qquad (3)\end{aligned}$$

where $\gamma_p = \gamma_m\sigma_p^2/\sigma_m^2$ and $\mathbf{e}_p = (\gamma_m/\sigma_m^2)\mathbf{VD}_p$ is the (unobserved) pricing error vector. This result is identical to that of Ashton and Tippett (1998, p.1333) except that they write $\mathbf{e}_p$ in terms of $N-2$ arbitrage portfolios. Clearly, the relation between $\boldsymbol{\beta}_p$ and $\mathbf{R}$ is less than perfect because of the presence of the pricing error $\mathbf{e}_p$.

The inability to observe the efficient market portfolio, $\mathbf{m}$, and the use of an inefficient proxy, $\mathbf{p}$, is equivalent to assessing the relation between $\boldsymbol{\beta}_p$ and $\mathbf{R}$. Thus, when the CAPM holds, the true model governing the relation between $\boldsymbol{\beta}_m$ and $\mathbf{R}$ is (1), while the true model governing the relation between $\boldsymbol{\beta}_p$ and $\mathbf{R}$ is (3). Both (1) and (3) are true models, but the reason why we are interested in (3) is our utilization of the inefficient index $\mathbf{p}$.

An objection might be that there exists an infinite number of 'true' models,

$$\mathbf{R} = r_{zm}\mathbf{1} + \gamma_p^*\boldsymbol{\beta}_p + \mathbf{e}_p^*, \qquad (4)$$

where γ_p^* and $\mathbf{e}_p^*$ are appropriately chosen. For example, $\gamma_p^* = \lambda\gamma_p\boldsymbol{\beta}_p$ and $\mathbf{e}_p^* = \mathbf{e}_p + (1-\lambda)\gamma_p\boldsymbol{\beta}_p$, for any arbitrary λ, are all deviations from (3). However, this is trivial since they all simplify algebraically to (3), which is the unique true relation between $\boldsymbol{\beta}_p$ and $\mathbf{R}$.

The point that any arbitrary model (4) reduces to (3) seems to have been missed by KS. These authors assume that expected returns are given by $\mathbf{R} = \mathbf{X}\theta + f$, where θ is some two-element vector, $\mathbf{X} = [\mathbf{1} \;\; \boldsymbol{\beta}_p]$, and f is an error vector. This model is identical to (4) with $\theta = [r_{zm} \quad \gamma_p^*]'$ and $\mathbf{X} = [\mathbf{1} \; \boldsymbol{\beta}_p]$. The true parameter relating $\boldsymbol{\beta}_p$ to $\mathbf{R}$ is γ_p and not an arbitrary γ_p^* (or, alternatively, θ).

A question that arises is whether there can be models where the true parameter relating $\boldsymbol{\beta}_p$ to $\mathbf{R}$ is different from γ_p. In other words, is it possible to have a model like (4) that reduces to (1) without passing through (3)? The answer is yes, but such a model would be inadmissible in a world of CAPM as it would imply that the index used in the test is not fully funded.

One interesting example is the deviation suggested by Ferguson and Shockley (2003), who assume that the true market portfolio consists of a weighted sum of the economy's debt claims ($\mathbf{D}$) and the economy's equity claims ($\mathbf{p}$)

$$\mathbf{m} = \lambda \mathbf{p} + (1 - \lambda)\mathbf{D}$$

where $0 < \lambda < 1$.

In our notation, $\mathbf{p}$ is an (observed) inefficient index and $\mathbf{D}$ is the omitted portfolio with $\mathbf{D}'\mathbf{1} = 1$. Replacing the value of $\mathbf{m}$ in equation (1) yields

$$\mathbf{R} = r_{zm}\mathbf{1} + \lambda \gamma_p \boldsymbol{\beta}_p + \frac{(1-\lambda)\sigma_D^2 \gamma_m}{\sigma_m^2}\boldsymbol{\beta}_D \tag{5}$$

Thus, the implicit assumption in this model is that the index used in computing betas is not fully funded ($\lambda\mathbf{p}$ rather than $\mathbf{p}$) and this violates the budget constraint in the CAPM optimization problem.

A simple conclusion from model (3) is that the true cross-sectional slope (γ_p) between expected returns and beta computed against any portfolio is always positive. However, although the true slope is non-zero, the empirical estimate of γ_p may not be so. Assuming we can measure $\boldsymbol{\beta}_p$ precisely, estimating r_{zm} and γ_p in model (3) suffers from the omitted regressor problem ($\mathbf{e}_p$) and will be biased unless $\mathbf{p} = \mathbf{m}$.

When $\mathbf{p}$ is on the upper half of the efficient frontier, $\mathbf{e}_p$ will be uncorrelated with $\boldsymbol{\beta}_p$ but correlated with the vector $\mathbf{1}$. Thus, the GLS/OLS estimators of r_{zm} are still biased, but they will both produce a perfect fit (as we shall see later). The reason is that betas computed from efficient portfolios are perfectly correlated.[2] So although the estimators are biased with respect to the true model (1), any beta computed against an efficient portfolio has a perfect linear relation with the cross-section of expected returns. I will exploit this divergence to derive a criterion for testing the CAPM.

When $\mathbf{p}$ is not on the efficient frontier, the empirical power of $\boldsymbol{\beta}_p$ to explain expected returns may or may not depend on the position of $\mathbf{p}$ relative to the efficient frontier. A significant finding by KS is that the choice of OLS versus GLS is critical. Although both methods yield biased estimates with respect to the true model, under OLS the empirical power of $\boldsymbol{\beta}_p$ suggested by the estimates (including the R^2) bears no relation to the proximity of portfolio $\mathbf{p}$ to the efficient frontier. The bias can be arbitrarily large or small.

On the other hand, under GLS the power of $\boldsymbol{\beta}_p$ to explain expected returns is strictly related to the position of portfolio $\mathbf{p}$ relative to the efficient frontier. The bias in both parameters is known exactly. Hence, one can use this GLS feature to assess the position of the proxy not only with respect to the efficient frontier but also with respect to the true market portfolio.

It is worth noting, however, that while KS have successfully explained the bias of the GLS estimator, they do not seem to have explained the behaviour of the OLS bias satisfactorily. Appendix 1 revisits KS analysis, with a view to providing a better explanation of the OLS bias.

4. Terminology: "Correctly Specified" and "True" Model

[2] If both $\mathbf{m}$ and $\mathbf{p}$ are efficient, $\boldsymbol{\beta}_p$ and $\boldsymbol{\beta}_m$ will be perfectly correlated because they are both perfectly correlated with $\mathbf{R}$.

Before discussing the proposed tests it is useful to remove potential ambiguity that can arise from the use of the terms 'correctly specified model' and 'true model'. The latter is related to the real world, so when we say that the CAPM is true (or, alternatively, that the CAPM holds), we simply mean that the data are actually generated by a CAPM model. The former is related to the empirical world. So when we say that the (empirical) model is correctly specified, we simply mean that the empirical model's specification coincides with the real data generating process.

In our context, a true expected return generating process can be written in the following general form $\mathbf{R} = r_{zm}\mathbf{1} + \gamma_m \boldsymbol{\beta}_m + \mathbf{d}$. This model allows for any cross sectional specification, such as APT or other multifactor models, through the vector $\mathbf{d}$. When $\mathbf{d} = \mathbf{0}$ I say that the CAPM holds, or the CAPM is true. This simply means that the assumption that 'the CAPM is the process that has generated the observed expected returns' is correct. An empirical CAPM model of the form $\mathbf{R} = r_{zm}\mathbf{1} + \gamma_m \boldsymbol{\beta}_m + \varepsilon$, where ε is the usual disturbance term,[3] will be said to be correctly specified. But this only happens when we observe the market portfolio (i.e. when we use $\boldsymbol{\beta}_m$ as a regressor).

Incorrect specification, or misspecification, takes place for two reasons. Either we wrongly assume that the observed portfolio is the market portfolio ($\mathbf{e}_p \neq \mathbf{0}$); and/or we wrongly assume that the CAPM holds ($\mathbf{d} \neq \mathbf{0}$). Thus, the (possibly non-CAPM) true model will be $\mathbf{R} = r_{zm}\mathbf{1} + \gamma_p \boldsymbol{\beta}_p + \mathbf{e}_p + \mathbf{d}$, whereas the empirical model will be $\mathbf{R} = r_{zm}\mathbf{1} + \gamma_p \boldsymbol{\beta}_p + \varepsilon$, but the disturbance term will no longer have zero mean or be uncorrelated with the regressor $\boldsymbol{\beta}_p$ (since $\varepsilon = \mathbf{e}_p + \mathbf{d}$).

[3] This term is usually assumed to have zero mean and to be uncorrelated with regressor(s). In this paper I assume away statistical or measurement errors, so in this case of correct specification, $\varepsilon = \mathbf{0}$.

Any inference carried out on the CAPM therefore crucially depends on whether the assumptions that the 'CAPM is true' and that 'the portfolio used to compute beta is the true market portfolio' are correct. If both assumptions are correct, then we would only have to contend with small sample or measurement errors (which are assumed away in this paper). However, when one or both of these assumptions are incorrect, then we would have to contend with the meaningfulness of the empirical estimation and inference.

In the next section, I will show that the CAPM can be tested in the simpler case where the assumption that the CAPM holds is correct ($\mathbf{d} = \mathbf{0}$), but the only potential problem is the possibility that the index is not the market portfolio ($\mathbf{e}_p \neq \mathbf{0}$). However, the inference will generally be invalid if the CAPM assumption is violated. Fortunately, there is a way to ensuring that $\mathbf{d} = \mathbf{0}$ before proceeding with the formal test of the CAPM. This is discussed in Section 6.

5. Testing the CAPM under the Assumption that the CAPM Holds

In this section I extend the results of KS, who show that the GLS method provides a useful statistic, namely the GLS R^2, that can be used to assess the efficiency of a given portfolio. I emphasise the point that KS build their results under the assumption that the CAPM is true. I will use the same assumption in this section to (i) show that the efficiency of a portfolio is a necessary but not sufficient condition for testing the CAPM; and (ii) derive a necessary and sufficient condition for testing the CAPM.

I start by showing the usefulness of GLS. The GLS bias, derived by KS (p.167), is given by

$$\begin{pmatrix} \hat{r}_{zm}(gls) \\ \hat{\gamma}_m(gls) \end{pmatrix} = \begin{pmatrix} r_{zm} \\ \gamma_m \end{pmatrix} + \left(1 - \psi_p\right) \begin{pmatrix} \gamma_m \frac{\sigma_g^2}{\sigma_m^2} \\ -\gamma_m \end{pmatrix}$$

where $\psi_p = (r_p - r_g)/(r_m - r_g)$, and r_g is the expected return of the Global Minimum Variance Portfolio (GMVP). For inefficient portfolios above the GMVP ($r_p > r_g$), the intercept is always overestimated and the slope is always underestimated. Also, for $r_p > r_g$, both GLS estimates will be positive. The slope bias is completely determined by portfolio **p**'s position in mean-variance space, but the intercept bias is only partially determined by ψ_p. Thus, for a given portfolio, different asset universes will produce identical slope biases but different intercept biases.

As KS argue, the OLS estimator offers no clues as to where portfolio **p** is situated relative to the efficient frontier. Relying on OLS can therefore be misleading. On the other hand, the GLS result is strictly related to the position of the inefficient portfolio in mean-variance space. The question then is how useful is GLS in testing the CAPM? The answer is that the raw quantities obtained from GLS are only useful for testing the efficiency of a given index. To see this, note that in the absence of sampling error, the GLS output contains three quantities: the intercept, $r_{zm} + (1 - \psi_p)(r_g - r_{zm})$, the slope $\psi_p(r_m - r_{zm})$, and the GLS R^2 given by[4]

$$\psi_p^2 = R^2_{gls} = 1 - \frac{(\mathbf{R} - \mathbf{X}\boldsymbol{\Gamma}_{gls})'\mathbf{V}^{-1}(\mathbf{R} - \mathbf{X}\boldsymbol{\Gamma}_{gls})}{(\mathbf{R} - \mathbf{1}\overline{\mu})'\mathbf{V}^{-1}(\mathbf{R} - \mathbf{1}\overline{\mu})}$$

where $\overline{\mu} = (\mathbf{R}'\mathbf{V}^{-1}\mathbf{1})/(\mathbf{1}'\mathbf{V}^{-1}\mathbf{1})$.

All three quantities depend on the measure of the relative efficiency of portfolio **p**, ψ_p. Although the intercept is uninformative, the sign of the GLS slope tells us the position of the index relative to the GMVP. A positive value of the slope indicates that the index is above the GMVP and vice versa. The R^2_{gls} informs us how close the index is to the frontier. Thus, only by combining the information contained in the slope and the R^2_{gls} can we determine the exact position of the inefficient index.

[4]The proof that $\psi_p^2 = R^2_{gls}$ can be found in KS, page 180.

However, determining the efficiency of a given index is not sufficient for testing the CAPM. Suppose the slope was positive and $\psi_p^2 = 1$. This would only tell us that portfolio **p** is efficient, but this would not be sufficient evidence in favour of the CAPM. In other words, the efficiency of an index is a necessary but insufficient condition for the CAPM to hold.

Although any efficient portfolio will provide perfect linearity and hence perfect fit, the CAPM requires that the slope equals the market expected return in excess of the zero beta rate, and that the intercept equals the zero beta rate. However, if an index, say **q**, is efficient the GLS regression would give two positive numbers

$$\mathbf{\Gamma}_{gls}^{q} = \begin{pmatrix} \hat{r}_{zq}(gls) \\ \hat{\gamma}_q(gls) \end{pmatrix} = \begin{pmatrix} r_{zq} \\ r_q - r_{zq} \end{pmatrix}$$

If the index is the market portfolio we would also be given two positive numbers

$$\mathbf{\Gamma}_{gls}^{m} = \begin{pmatrix} \hat{r}_{zm}(gls) \\ \hat{\gamma}_m(gls) \end{pmatrix} = \begin{pmatrix} r_{zm} \\ r_m - r_{zm} \end{pmatrix}$$

Both **q** and **m** will satisfy $\psi_m^2 = \psi_q^2 = 1$. Telling the difference between **q** and **m** is the key to the testability of the CAPM. Using the relationship betwen the GMVP and any efficient portfolio, we have

$$r_g = r_{zq} + (r_q - r_{zq})\tfrac{\sigma_g^2}{\sigma_q^2}$$
$$r_g = r_{zm} + (r_m - r_{zm})\tfrac{\sigma_g^2}{\sigma_m^2}$$

Combining these two equations gives

$$r_q = r_{zq} + \frac{\sigma_q^2}{\sigma_g^2}(r_{zm} - r_{zq}) + \frac{\sigma_q^2}{\sigma_m^2}(r_m - r_{zm}) \tag{6}$$

From (3)

$$
\begin{aligned}
\mathbf{R} &= r_{zm}\mathbf{1}+\frac{\sigma_q^2\gamma_m}{\sigma_m^2}\boldsymbol{\beta}_q+\frac{\gamma_m}{\sigma_m^2}\ (\mathbf{Vm}-\mathbf{Vq}) \\
&= (r_{zm}+\alpha_q\frac{\gamma_m}{\sigma_m^2})\mathbf{1}+\frac{\sigma_q^2\gamma_m}{\sigma_m^2}\boldsymbol{\beta}_q \\
&= r_{zq}\mathbf{1}+(r_q-r_{zq})\boldsymbol{\beta}_q \qquad (7)
\end{aligned}
$$

where α_q is a constant and results from the fact that $\boldsymbol{\beta}_q$ and $\boldsymbol{\beta}_m$ are perfectly correlated. Suppose we estimate a model using the efficient portfolio $\mathbf{q}$. Using (7) we have

$$\hat{r}_{zq}(gls)+\hat{\gamma}_q(gls)=r_{zq}+\tfrac{\sigma_q^2}{\sigma_m^2}(r_m-r_{zm})$$

Subtracting (6) from above yields

$$\hat{r}_{zq}(gls)+\hat{\gamma}_q(gls)-r_q=\tfrac{\sigma_q^2}{\sigma_g^2}(r_{zm}-r_{zq})$$

which only equals zero if $\mathbf{q}=\mathbf{m}$.

Thus, GLS can help us not only to determine the efficiency of a given index, but also to test the CAPM. It exploits the fact that under the null ($\mathbf{q}=\mathbf{m}$)

$$\hat{r}_{zq}(gls)+\hat{\gamma}_q(gls)-r_q=0$$

So, one can test whether the sum of slope and intercept are equal to the expected index return. Note that when the index is efficient, OLS and GLS are identical in the absence of sampling error. Thus, in principle, both methods could be useful in testing the CAPM. However, empirical investigators use sample estimates of returns, betas and the covariance matrix. That is, empiricists use feasible GLS rather than GLS. So, as Shanken (1992) and KS caution, there is a potential concern "about the GLS estimator in finite samples" (KS, p.170). Nevertheless, in practice the choice between OLS and GLS will depend on considerations other than the population moments. Although it is beyond the scope of this paper to discuss these practical choices, it should be noted that most econometric textbooks describe to some length the conditions under which some estimators might be preferable to others.

6. Omitted Factors

The previous analysis was based on the assumption that the SLB CAPM holds. The only source of potential problems considered so far has been the use of an inefficient index. This follows the tradition of RR, KS and Ferguson and Shockley (2003) who do not consider the case where the CAPM does not hold. However, since my results support the testability of the CAPM when it is true, I need to assess whether the CAPM can also be rejected when it is false.

Appendix 2 shows that there is no loss of generality if we focused on vertical departures from the market portfolio. Thus, let $\mathbf{X} = [\mathbf{1} \quad \boldsymbol{\beta}_p]$ and $\boldsymbol{\Gamma} = [r_{zm} \quad \gamma_m]'$, and consider the general (possibly non-CAPM) model

$$\mathbf{R} = \mathbf{X}\boldsymbol{\Gamma} + \mathbf{e}_p + \mathbf{d}$$

where $\mathbf{d}$ is not proportional to the vector $\mathbf{1}$, and $\mathbf{e_p} \neq -\mathbf{d}$.[5] This is the case of both omitting relevant factors and using a possibly inefficient index to compute betas. When $\mathbf{d} \neq \mathbf{0}$, the previous results on GLS bias no longer hold, even if $\mathbf{e}_p = \mathbf{0}$.[6] Thus, we are now faced with two potential sources of bias, and we cannot identify whether the source of bias is due to the inefficiency of the index or due to misspecification.

The bias for the OLS and GLS estimators is given by, respectively

[5] $\mathbf{e}_p + \mathbf{d} = \mathbf{0}$ implies that the CAPM holds and that $\mathbf{p}$ is efficient.

[6] However, when $\mathbf{d}$ is orthogonal to $\mathbf{X}$ the OLS will be unbiased, and when $\mathbf{d}$ is orthogonal to $\mathbf{XV}^{-1}$ the GLS will be unbiased.

$$\mathbf{\Gamma}_{ols} = \mathbf{\Gamma} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'(\mathbf{e}_p + \mathbf{d})$$
$$\mathbf{\Gamma}_{gls} = \mathbf{\Gamma} + (\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}(\mathbf{e}_p + \mathbf{d})$$

It is clear that the bias would only vanish when $\mathbf{e}_p = \mathbf{0}$ (that is, the index used to compute beta is the market portfolio) and $\mathbf{d}$ is orthogonal to $\mathbf{X}$ (for the OLS case), or $\mathbf{V}^{-1}\mathbf{X}$ (for the GLS case). Note that the bias will remain if the index was efficient but not equal to the market portfolio (that is, if $\mathbf{e}_p$ was proportional to the vector **1).** Thus, omitting factor sensitivities may worsen the spuriousness problem because there are now two sources of bias. More importantly, the GLS bias is no longer related to the position of the index relative to the efficient frontier. This is simply because $\mathbf{d}$ may not be related to $\mathbf{V}^{-1}$ and $\mathbf{X}$ in the same way as $\mathbf{e}_p$.

As a result, R^2_{gls} loses its strict relation with the index position relative to the efficient frontier and, like R^2_{ols}, may behave erratically. But this is not the only consequence of admitting potential misspecification. In KS, and under the assumption that the CAPM is true generating process, $R^2_{ols} = 1$ or $R^2_{gls} = 1$ can only happen when the index is efficient. A surprising result here is that the R^2 can attain its maximal value even when the index is inefficient. This is because there are now two sources of bias, $\mathbf{e}_p$ and $\mathbf{d}$, which can combine to give $R^2 = 1$.

To see this, note that from the definition of the R^2_{ols} (Appendix 1) and R^2_{gls}, the vectors of interest are, respectively

$$Q_{ols} = \mathbf{R} - \mathbf{X}\mathbf{\Gamma}_{ols} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\,(\mathbf{e}_p + \mathbf{d})$$
$$Q_{gls} = \mathbf{R} - \mathbf{X}\mathbf{\Gamma}_{gls} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}]\,(\mathbf{e}_p + \mathbf{d})$$

These are the pricing errors obtained from the OLS and GLS estimation, respectively. Because both R^2 involve quadratic terms, $R^2_{ols} = 1$ and $R^2_{gls} = 1$ if and only if $Q_{ols} = \mathbf{0}$ and $Q_{gls} = \mathbf{0}$, respectively. This is obviously satisfied when (i) $\mathbf{e}_p = \mathbf{0}$, or $\mathbf{e}_p$ is proportional to the vector $\mathbf{1}$, and (ii) $\mathbf{d} = \mathbf{0}$, that is, the index used to compute beta is efficient and there are no omitted factor sensitivities (i.e. the CAPM is true).

However, there is a possibility that $R^2 = 1$ even when the above conditions are not met. For the OLS, when $\mathbf{d}$ is orthogonal to $\mathbf{X}$ (i.e. $\mathbf{X}'\mathbf{d} = 0$) the pricing error vector simplifies to

$$Q_{ols} = \left[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\right] \mathbf{e}_p + \mathbf{d}$$

Thus, $\mathbf{d} = \left[\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' - I\right] \mathbf{e}_p$ implies $Q_{ols} = \mathbf{0}$.[7]

Similarly, when $\mathbf{d}$ is orthogonal to $\mathbf{X}\mathbf{V}^{-1}$ (i.e. $\mathbf{X}'\mathbf{V}^{-1}\mathbf{d} = 0$) the pricing error vector simplifies to

$$Q_{gls} = \left[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}\right] \mathbf{e}_p + \mathbf{d}$$

So, $\mathbf{d} = \left[\mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1} - I\right] \mathbf{e}_p$ implies $Q_{gls} = 0$.

Thus, one could unawarily use both an inefficient index while omitting relevant variables from the empirical model and still obtain a perfect OLS or GLS fit. Fortunately, we cannot obtain a perfect fit for both, for even in the unlikely event that $\mathbf{d}$ was orthogonal to both matrices (i.e. $\mathbf{X}'\mathbf{d} = \mathbf{X}'\mathbf{V}^{-1}\mathbf{d} = \mathbf{0}$) it would not be able to satisfy $Q_{ols} = Q_{gls} = 0$. The reason is that[8]

$$\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}' \neq \mathbf{X}(\mathbf{X}'\mathbf{V}^{-1}\mathbf{X})^{-1}\mathbf{X}'\mathbf{V}^{-1}$$

Thus, when the CAPM is not the true generating process ($\mathbf{d} \neq \mathbf{0}$), using an efficient index actually ensures that both OLS and GLS R^2 will be less

[7] We exclude the case where $\mathbf{e}_p$ is a linear combination of the columns of $\mathbf{X}$, because that would imply $\mathbf{d} = 0$. This also applies to the GLS case below.

[8] Except for the trivial case where the covariance matrix is proportional to the identity matrix.

than 1, regardless of whether or not $\mathbf{d}$ is orthogonal to $\mathbf{X}$ or $\mathbf{XV}^{-1}$. This highlights the danger of carrying out inference under the assumption that the CAPM holds. Were we to do that in the presence of an omitted sensitivity, we could conclude, wrongly, that the index is inefficient on the basis of the GLS R^2 as suggested by KS. Unexpectedly, when the index is inefficient there is a possibility that one of the R^2 will equal 1. Again, were we to believe one of the R^2 we would wrongly conclude that the index is efficient, when in fact it is not. Fortunately, there is a unique case where we can ensure that the index is efficient and that the CAPM is the true model, namely when both OLS and GLS R^2 attain the maximum value of 1.[9]

Thus, one major conclusion emerges: the CAPM is testable even under potential misspecification (inefficient index and omitted factor loadings). It is well known that omitted variables create inference problems because we do not know whether the estimates are biased or not. Econometricians are aware of this problem and generally tolerate it, especially when the omitted regressor is orthogonal to, or at least uncorrelated with, the right hand side variables. Most empirical models in economics accept the less than perfect relation between regressor(s) and regressee, and often focus on obtaining unbiased estimates. One further advantage with economic models is that regressors are observed (or assumed to be observed). Thus, an R^2 that is smaller than 1 is not a cause for concern. In CAPM language, this is equivalent to saying that (i) the CAPM holds; (ii) the CAPM predicts an imperfect relation between beta and expected returns; and (iii) the market portfolio is observable. Were we to have these facilities, we would then only be interested in the market risk premium, hoping that omitted factors sensitivities were

[9] Note that, here, efficient means that the portfolio is on the minimum variance frontier. So maximum coefficient of determinations can also obtain from a perfectly inefficient portfolio.

uncorrelated with the market beta. The R^2 would be of no consequence.

Unfortunately, when testing the CAPM the econometrician does not benefit from such facilities. Even with no measurement errors, the econometrician still needs to show that the observed index is efficient, that it is the market portfolio, and that there are no omitted factor sensitivities (i.e. the CAPM holds). RR argue that the first requirement cannot be attained using OLS even if we assumed that the CAPM holds. KS show that index efficiency can be assessed via GLS when the CAPM holds, but do not explore the property of GLS in a non-CAPM world. Understandably, given the pessimism against the testability of the CAPM, assessing whether the index is the market portfolio and whether there are no omitted variables has not been dealt with previously. In this paper, I offer a positive answer to the sceptics: the CAPM is testable. Of course, like RR and KS I do assume measurement errors away. What happens when there are measurement errors is a different, and certainly more challenging, undertaking. The aim of this paper is much more modest. I have simply attempted to establish a theoretical principle, namely the CAPM is testable.

7. A Simulation Experiment

Although we know the exact behaviour of the GLS slope, we are far less informed about the OLS coefficients. All we know is that they are biased and that the bias is arbitrary and is not related to the proximity of the index to the efficient frontier. Since it is not possible to assess the scale and sign of the OLS bias theoretically, it should be useful to infer the behaviour of the OLS estimates via a simulation experiment.

The simulation is carried out in two main steps. The first step involves generating a mean-variance frontier, with a universe of $N = 10$ asset returns and their associated variances and covariances. In the second step I take the mean-variance universe and randomly generate deviations from the market return, and calculate the OLS and GLS estimates of parameters and R^2.

First, to generate a mean-variance frontier, a positive definite variance-covariance matrix, $\mathbf{V}$, is randomly drawn such that the GMVP's standard deviation lies between 2% and 6%. Given $\mathbf{V}$, I randomly draw a set of market weights, $\mathbf{m}$, such that the market portfolio's standard deviation is less than 30%. More specifically, I first obtain beta, $\boldsymbol{\beta}_m = \mathbf{Vm}/\mathbf{m}'\mathbf{Vm}$, and then obtain a set of compatible expected returns, $\mathbf{R} = \theta_1 + \theta_2\boldsymbol{\beta}_m$. The zero beta rate θ_1 is drawn from a uniform distribution with a range of 0.4% to 0.6%. The excess market return θ_2 is drawn from a uniform distribution with a range of 0.65% to 0.85%. The market return and variance are then obtained as $r_m = \mathbf{m}'\mathbf{R}$, and $\sigma_m^2 = \mathbf{m}'\mathbf{Vm}$. I redraw $\mathbf{m}$ and repeat the procedure until σ_m is less than 30%.[10]

Next I generate random deviations from the market portfolio to obtain an inefficient portfolio $\mathbf{p}$. Since it is difficult to obtain vertical deviations from $\mathbf{m}$, I generate the efficient portfolio, $\mathbf{q}$, that has the same variance as $\mathbf{p}$. The comparison is then made between $\mathbf{p}$ and $\mathbf{q}$. Thus, for each replication the true model is $\mathbf{R} = r_{zq} + \gamma_q\boldsymbol{\beta}_q$, where $\gamma_q = r_q - r_{zq}$. The OLS and GLS estimates are obtained using $\boldsymbol{\beta}_p$ instead of $\boldsymbol{\beta}_q$. The coefficients are then compared with the true values r_{zq} and γ_q. I also calculate ψ_p and the coefficients of determination for each replication.

To minimise the chance that the results are driven by an arbitrary choice of the efficient frontier, I maximise the number of asset universes used in

[10] To give an idea about the kind of frontiers I use in this simulation, I generated 5000 frontiers (5 runs of 1000 replications) using both positive and unconstrained weighting schemes. The average market return was 1.25% with a range of (1.05%, 1.44%); the average market standard deviation was 16% with a range of (3%, 30%); the GMVP return averaged 0.60% with a range of (0.41%, 1.10%). Finally, the GMVP standard deviation averaged 3.10% with a range of (2.00%, 5.90%). Asset standard deviations ranged from a minimum of 6% to a maximum of 27%.

the simulation. I generate 1000 different frontiers, and for each of these I generate 1000 deviations from the market portfolio. These deviations are computed such that a wide range of portfolio inefficiency is produced. This gives 1 million replications, producing GLS and OLS biases.

A. Comparing the Slopes.

The basis for comparing the slopes is the relative bias, which is given by $\frac{\hat{\gamma}_q - \gamma_q}{\gamma_q}$. However, to get an idea about how distant the estimate is from the true value I use a slight modification, $1 + \frac{\hat{\gamma}_q - \gamma_q}{\gamma_q}$. This shows the bias as a proportion of the true slope. For example, a value of 0.40 means that $\hat{\gamma}_q = 0.4\gamma_q$, that is, the OLS coefficient underestimates the true parameter by 60%.

The GLS and OLS biases will behave according to, respectively,

$$\begin{aligned} 1 + \frac{\hat{\gamma}_q(gls) - \gamma_q}{\gamma_q} &= \psi_p \\ 1 + \frac{\hat{\gamma}_q(ols) - \gamma_q}{\gamma_q} &= \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_q)}{v\hat{a}r(\boldsymbol{\beta}_p)} \end{aligned}$$

Figure 1 shows these biases for two cases, each of which depends on whether or not a constraint of positive portfolio weights is imposed. These are (a) constrained **m** and **p**, and (b) unconstrained **m** and **p**.[11] This ensures that the results are not driven by short positions (RR, 1994; Grauer, 1999).

As expected, the GLS bias is exact in ψ_p. The constrained and unconstrained cases are very similar and confirm the arbitrariness of the OLS bias. The spread around the true slope is large even for highly efficient indexes. Zero estimated slopes are not expected to be uncommon, and for indexes that are roughly less than 80% efficient we should even expect negative OLS slopes. Yet, large estimates cannot be ruled out.

[11] The cases where only one of the two portfolios is constrained were very similar throughout and are therefore not reported.

B. Comparing the Intercepts.

The GLS intercept has an exact relation with ψ_p, but is not exclusively determined by it. However, it is possible to derive a bias criterion such that it is exactly linear in ψ_p. Using the GLS bias result, and equation (A66), p.180 of KS, I obtain

$$\frac{\hat{r}_{zq}(gls)-r_{zq}}{r_g-r_{zq}} = 1-\psi_p$$

Unfortunately, this cannot be compared with a similar OLS criterion. Using the OLS bias result gives

$$\frac{\hat{r}_{zq}(ols)-r_{zq}}{r_g-r_{zq}} = \frac{r_q-r_{zq}}{r_g-r_{zq}}\left(\bar{\beta}_q-\bar{\beta}_p\frac{c\hat{o}v(\boldsymbol{\beta}_p,\boldsymbol{\beta}_q)}{v\hat{a}r(\boldsymbol{\beta}_p)}\right)$$

So, although the GLS intercept will not depend on the frontier, the OLS criterion will depend on three frontier parameters. The OLS bias would appear to be arbitrary by construction. However, the above OLS criterion simplifies to a relative measure that will not depend on the frontier

$$\frac{\hat{r}_{zq}(ols)-r_{zq}}{r_q-r_{zq}} = \left(\bar{\beta}_q-\bar{\beta}_p\frac{c\hat{o}v(\boldsymbol{\beta}_p,\boldsymbol{\beta}_q)}{v\hat{a}r(\boldsymbol{\beta}_p)}\right)$$

This gives a bias relative to the efficient index return rather than the GMVP return. The OLS and GLS bias criteria are therefore not directly comparable. Nervertheless, there is a clear difference. The weighted GLS bias should be exactly linear with the efficiency of the index. On the other hand, the weighted OLS bias is expected to be unrelated to the efficiency of the index.

Figure 2(a) plots the relative OLS bias for the constrained case (positive **m** and **p**). For highly inefficient indexes, the OLS method seems to always overestimate the intercept. However, for more efficient indexes the bias can go in either direction. In particular, a negative value for the OLS estimate,

$\hat{r}_{zq}(ols)$, cannot be ruled out. In the above experiment, around 600 intercepts (out of the million replications) were negative.

Figure 2(b) shows a plot of the OLS relative bias when the market and inefficient portfolios are allowed to have short positions. The relation between relative efficiency and intercept seems flatter than that of the constrained case, but the bias is still arbitrary.

In both constrained and unconstrained cases there is an obvious tendency for the OLS to over-estimate the intercept. The 1st and 99th percentiles suggest that the vast majority of cases have positive bias. As expected, the GLS relative bias is exactly linear and is therefore not reported. Clearly, the GLS estimate of the intercept is always biased upwards and hence always positive.

C. The R-Squared.

The R^2 results for the experiment are plotted in Figure 3. Both constrained (positive $\mathbf{m}$ and $\mathbf{p}$) and unconstrained cases are shown. The GLS shows the expected exact curve, but the OLS bias appears to depend on the constraint. While the unconstrained bias is smoother and fairly symmetric, the constrained percentiles are tilted to the right. However, both cases give the same message. The OLS varies widely with relative portfolio efficiency, ψ_p. If we take the extremes, zero R^2_{ols} are very common and can be found for indexes that are more than 95% efficient in the constrained case. Large values for the R^2_{ols} are also common at all levels of inefficiency. In general it is clear that the OLS does not provide guidance as to the efficiency of a given portfolio.

D. *Nearly Efficient Portfolios.*

As mentioned earlier, one advantage of the bias derived in this paper is that it allows for a bridge between the efficient and inefficient cases. From the bias formulas we expect the slope bias and R^2_{ols} to converge to zero and one respectively, as the index gets arbitrarily close to the upper half of the efficient frontier. This is shown in Figure 4, where the OLS slope bias and R^2_{ols} for the 4000 most efficient simulated portfolios are plotted. The slope bias is still substantial for portfolios that are 99.3% efficient. However, we can also see that the OLS slope estimates get closer to the true slope as the portfolio efficiency gets arbitrarily close to 100%, but we would not be comfortable with portfolios that are less than 99.99% efficient. For example, a 99.8% efficient portfolio can produce a slope that is anywhere between less than 90% and almost 110% the true slope.[12] This clearly casts doubt on the usefulness of OLS even with nearly efficient indexes. The R^2_{ols} shows a similar pattern. Figure 4 shows that R^2_{ols} of less than 0.96 can still be obtained for indexes that are between 99.7% and 99.8% efficient. Again, for very nearly efficient portfolios, the R^2_{ols} is very close to its maximum value.

[Insert Figures 1-4 around here]

E. *Omitted Factors.*

The above simulations are based on the assumption that the true data generating process is the CAPM ($\mathbf{d} = \mathbf{0}$). The R^2_{gls} is exactly related to the efficiency of the index as expected, while the R^2_{ols} seems to converge to its maximum value as the index gets arbitrarily close to the efficient frontier. With potential omitted factors, the interest lies in the behaviour of both coefficients of determination. To show this, expected returns are generated as before, except that one additional term is added to the CAPM, $\mathbf{R} = r_{zq} + \gamma_q \boldsymbol{\beta}_q + \mathbf{d}$. I generate omitted sensitivities, $\mathbf{d} = \alpha \mathbf{d}^*$, by taking seven

[12] These values are true for this particular experiment. Different simulations may produce wider bounds.

different values for the scale, α, starting at zero (CAPM) and then increasing to 0.001, 0.01, 0.05, 0.1, 0.5 and 1. For each value of α a random vector $\mathbf{d}^*$ is drawn from a uniform distribution with a range of 0% to 1%, and then the whole simulation procedure is repeated (that is, 1000 different frontiers by 1000 deviations from the market portfolio). This gives one million replication for each of the seven scales, α.

The results are shown in Figure 5. The first pair of plots show the 1st and 99th percentiles for the R^2_{ols} for each of the seven scales of omission. Because OLS is highly sensitive to repackaging, the first percentile is as low as 29.3% for indexes that are 90% efficient. The addition of omitted variable has some effect, especially for indexes below 96% efficiency. Table 1 shows the values of the R^2_{ols} percentiles. In general the R^2_{ols} tends to decrease with increasing importance of omitted sensitivities. More importantly, the R^2_{ols} can reach the maximum value even for indexes that are not perfectly efficient. This is shown under $\alpha = 0.05$ where the 99th percentile equals exactly 1 for indexes whose efficiency is between 99% and 99.5%. Although the general trend is for the R^2_{ols} to increase with increasing index efficiency, the 1st percentile remains relatively low (between 97.2% and 97.9%). We also get a glimpse on the effect of more important omission at the highest scales ($\alpha = 0.5$ and 1). All 99th percentiles decrease relative to other lower scales. Unfortunately, I have been unable to obtain highly efficient indexes with higher scales and I leave this intersting behaviour for a future investigation.

The GLS statistics are far less sensitive to repackaging, and show a much smoother relation with the index efficiency and the omission scale. Table 1 reveals that the difference between the 1st and 99th percentiles is around 1% for all cases. The increase in the first five scale values does not seem to have any significant impact (the values are not identical but are too small to be shown at the third decimal place). However, the two highest scales have a clear impact on the 1st and 99th percentiles for the two highest levels of efficiency. For up to 99.5% index efficiency, the 99th percentile goes down from 98.9% to 98.4% when the omission scale increases from 0.5 to 1. Similarly, when the omission scale increases from 0.1 to 0.5, the 99th percentile decreases from 99.9% to 99.4% for the highest efficiency indexes. As before, the increase in index efficiency increases the R^2_{gls} overall.

To sum up, the impact of omitted sensitivities on the coefficients of determination depends on how important these sensitivities are to expected returns. For low importance, the GLS is generally insensitive and will depend mainly on the efficiency of the index. The OLS is highly sensitive and can help reject the CAPM even for very highly efficient indexes. As these omissions become more and more important, the limited results of the simulation indicate that both R^2_{ols} and R^2_{gls} will start to decrease sharply.

[Insert Figure 5 around here]

[Insert Table 1 around here]

8. Conclusions

Testing the CAPM has long been problematic. The use of an inefficient index is by no means its only limitation, but the possibility that the indexes used typically in empirical tests of the CAPM are inefficient undermines even the question of whether we are really testing the CAPM or a model other than the CAPM. The most challenging question to previous empirical studies of cross-sectional tests of the CAPM has been whether the CAPM is testable at all. Even assuming estimation problems away, one is still faced with two major uncertainties: the inefficiency of the market index and the possibility that the CAPM does not hold (omitted variables).

In this paper I propose a new way of testing the CAPM, even in the presence of orthogonal omitted factor sensitivities. I show that the CAPM is testable both when the CAPM holds and when it does not. The trick is to use both OLS and GLS. The only way that both R^2 can attain their maximum value is when we use an efficient index and there are no omitted variables. Once it is ascertained that the CAPM is true and that the index is efficient, one can then proceed to formally test the CAPM using an extension of KS results.

The results of this paper are based on the assumption that there are no sampling errors. This assumption helps to isolate the fundamental behaviour of the model at hand. In practice, though, the OLS and GLS estimates will be subject to additional uncertainty due to measurement errors in expected returns and betas. The literature has come up with many suggestions as to how to deal with the error in variable (beta) problem, and these seem to have been generally accepted. Also, the trade-off between OLS and GLS is well known in the econometric literature. There is less agreement, though, on how to deal with the question of "error in expected return" (Merton, 1980; Elton, 1999), and this remains a problem for all cross sectional tests of asset pricing models.

In addition, testing the CAPM now requires a pre-test involving R^2_{ols} and R^2_{gls}. The next step is, therefore, to derive the statistical properties of R^2_{ols} and R^2_{gls} under both the CAPM and the non-CAPM alternative. A recent paper by Kan, Robotti and Shanken (2013) provides some asymptotic results of the individual cross-sectional R^2. An interesting extension of their work would be to derive asymptotic properties of a test involving a combination of R^2_{ols} and R^2_{gls}. For example, the pre-test advocated in this paper would involve the joint null hypothesis: $R^2_{ols} = R^2_{gls} = 1$.

References

Ashton, D., and M. Tippett, 1998, Systematic risk and empirical research, *Journal of Business Finance and Accounting* 25, 1325-1355.

Black, F., 1972, Capital market equilibrium with restricted borrowing, *Journal of Business* 45, 444-455.

Elton, E.J., 1999, Expected Return, Realized Return, and Asset Pricing Tests, *Journal of Finance* 54, 1199-1220.

Ferguson, M.F., and R.L. Shockley, 2003, Equilibrium "Anomalies", *Journal of Finance* 58, 2549-2580.

Grauer, R.R., 1999, On the cross-sectional relation between expected returns, betas, and size, *Journal of Finance* 54, 773-789.

Grauer, R.R. and J.A. Janmaat, 2009, On the power of cross-sectional and multivariate tests of the CAPM, Journal of Banking & Finance, 33, 775–787.

Kan, R., C. Robotti and J. Shanken, 2013, Pricing Model Performance and the Two-Pass Cross-Sectional Regression Methodology, *Journal of Finance*, 68, 2617-2649.

Kandel, S., and R.F. Stambaugh, 1995, Portfolio inefficiency and the cross-section of expected returns, *Journal of Finance* 50, 157-184.

Lintner, J., 1965, The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets, *Review of Economics and Statistics* 47, 13-37.

Merton, R.C., 1980, On estimating the expected return on the market: An exploratory investigation, *Journal of Financial Economics* 8, 323-361.

Murtazashvili I., and N. Vozlyublennaia, 2012, The performance of cross-sectional regression tests of the CAPM with non-zero pricing errors, Journal of Banking & Finance 36, 1057–1066.

Roll, R., 1977, A critique of asset pricing theory's tests; Part 1: On past and potential testability of the theory, *Journal of Financial Economics* 4, 129-176.

Roll, R., 1980, Orthogonal portfolios, *Journal of Financial and Quantitative Analysis*, 15, 1005-1023.

Roll, R., and S.A. Ross, 1994, On the cross-sectional relation between expected returns and betas, *Journal of Finance*, 49, 101-121.

Shanken, J., 1992, On the estimation of beta-pricing models, *The Review of Financial Studies* 5,1-33.

Sharpe, W.F., 1964, Capital asset prices: A theory of market equilibrium under conditions of risk, *Journal of Finance* 9, 425-442.

Appendix 1. The KS Analysis: OLS versus GLS

KS analyse the bias in both OLS and GLS estimators when the betas are computed using an inefficient portfolio. However, they do not offer ways to directly compare the true OLS estimates, θ, with the true GLS estimates, ϕ. First, the OLS bias is based on an arbitrary parameter vector θ. They assume that expected returns are given by $\mathbf{R} = \mathbf{X}\theta + f$, where θ is some two-element vector, $\mathbf{X} = [\mathbf{1} \ \ \boldsymbol{\beta}_p]$, and f is an error vector. Although this is similar to equation (3), it is not related to the true process (1) as neither θ nor f are explicitly linked to the true mean-variance parameters. Thus, the OLS estimator is compared to an arbitrary 'true' parameter vector, rather than one that is implied by a theoretical model. While this is useful in demonstrating that the bias can be arbitrarily small or large, it does not allow us to compare it with the GLS estimator ϕ. In this latter case, KS show that the GLS estimator is related to the returns of the efficient portfolio with the same variance as portfolio $\mathbf{p}$, the zero beta portfolio, and the global minimum variance portfolio. But in both cases, there is a lack of a true theoretical model to compare the performance of these two estimation methods.

Relating the two estimation methods to the same reference model makes it possible to relate the bias to a single set of true parameters. It is worth noting also that the OLS bias I derive in the next subsection is valid for inefficient as well as efficient portfolios. In contrast, the bias derived in KS is only valid for strictly inefficient portfolios.

A. The OLS Bias.

Without loss of generality, I focus on inefficient portfolios that have the same variance as their efficient portfolio. In other words, I only focus on 'vertical' departures, $\mathbf{p}$, from an efficient portfolio, $\mathbf{m}$, such that $\sigma_p^2 = \sigma_m^2$.[13]

[13] Appendix 2 shows that the bias for inefficient indexes with different variances from

Let $\mathbf{X} = [\mathbf{1} \;\; \boldsymbol{\beta}_p]$ so that model (3) can be written as

$$\mathbf{R} = \mathbf{X\Gamma} + \mathbf{e}_p$$

where $\boldsymbol{\Gamma} = [r_{zm} \quad \gamma_m]'$, $\gamma_m = \gamma_p = r_m - r_{zm}$, and $\mathbf{e}_p = (\gamma_m/\sigma_m^2)(\mathbf{Vm} - \mathbf{Vp}) = \gamma_m(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p)$.

The general form of the OLS estimator is given by $\boldsymbol{\Gamma}_{ols} = \boldsymbol{\Gamma} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}_p$. Appendix 2 shows that this is equal to

$$\begin{pmatrix} \hat{r}_{zm}(ols) \\ \hat{\gamma}_m(ols) \end{pmatrix} = \begin{pmatrix} r_{zm} \\ \gamma_m \end{pmatrix} + \gamma_m \begin{pmatrix} \bar{\beta}_m - \bar{\beta}_p \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} \\ \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} - 1 \end{pmatrix}$$

were $\bar{\beta}$, $v\hat{a}r$ and $c\hat{o}v$ are the sample (cross-sectional) mean, variance and covariance respectively. The second term of the right-hand side is the bias and is generally only equal to zero when $\mathbf{p} = \mathbf{m}$.[14] However, the value of this bias can be arbitrarily small or large and does not depend on the degree of efficiency of portfolio $\mathbf{p}$.

There are at least two arguments why this bias may not be related to portfolio $\mathbf{p}$'s position in mean-variance space. First, there is an infinite number of portfolios that have the same variance and mean as $\mathbf{p}$. Thus, there are always some portfolios that can produce betas yielding arbitrary values for the bias. Second, we can use the repackaging argument. Since repackaging does not alter portfolio $\mathbf{p}$'s location in mean-variance space, repackaging the assets produces the same 'inefficient' betas, but different market betas and hence different average market betas and covariance/variance ratios. Because the bias is affected by repackaging we cannot systematically quantify the behaviour of the bias.

that of the market is unaffected.

[14] Strictly speaking, the bias should be the expectation of the second term of the right-hand side. But since all betas are assumed to be measured without error, the expectation of the term equals the term itself. This applies to all biases derived in this paper.

More precisely, KS show that one can repackage the assets (i.e. change their means and variances) without changing the inefficient portfolio's position or its betas. In other words, there exists a matrix A such that $A\mathbf{1} = \mathbf{1}$, and $A\boldsymbol{\beta}_p = \boldsymbol{\beta}_p$ (equations A19 and A20, p.174). However, for the efficient portfolio, $A\boldsymbol{\beta}_m \neq \boldsymbol{\beta}_m$.[15]

Multiplying both sides of the identity $\mathbf{m} = \mathbf{p} + \mathbf{D}_p$ by $\mathbf{V}$, and dividing by $\sigma_m^2 = \sigma_p^2$ we obtain $\boldsymbol{\beta}_m = \boldsymbol{\beta}_p + \mathbf{u}_p$, where $\mathbf{u}_p = \mathbf{V}\mathbf{D}_p/\sigma_m^2$. Multiplying throughout by the repackaging matrix gives $A\boldsymbol{\beta}_m = \boldsymbol{\beta}_p + A\mathbf{u}_p$. Thus, the relationship between the two betas depends on repackaging, which produces different market betas for each repackaging. Different market betas means different $\bar{\beta}_m$ and $c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)$, which leads to the arbitrariness of the OLS bias.

B. The OLS R-Squared.

Propositions 1 and 2 (page 160 and 163) of KS state that, if the market index is inefficient, the OLS R^2 can have essentially any value ($0 < R^2_{ols} < 1$). I confirm their result but use simpler algebra.

The OLS R^2 is given by

$$R^2_{ols} = 1 - \frac{(\mathbf{R}-\mathbf{X}\boldsymbol{\Gamma}_{ols})'(\mathbf{R}-\mathbf{X}\boldsymbol{\Gamma}_{ols})}{(\mathbf{R}-\frac{\mathbf{1}'\mathbf{R}}{N}\mathbf{1})'(\mathbf{R}-\frac{\mathbf{1}'\mathbf{R}}{N}\mathbf{1})}$$

Thus, the behaviour of the R^2_{ols} depends on the term $\mathbf{R} - \mathbf{X}\boldsymbol{\Gamma}_{ols} = [\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\,\mathbf{e}_p$. The R^2_{ols} will always be less than one, unless $\mathbf{p}$ is efficient. This is simply because $[\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}']\mathbf{e}_p$ cannot equal a vector of zero since $\mathbf{e}_p$ is not perfectly correlated with any of the columns of $\mathbf{X}$.[16]

To see this more precisely, rewrite the OLS bias as

[15] To see this, note that an efficient portfolio produces an exact relation between beta and expected returns, $\mathbf{R} = r_{zm}\mathbf{1} + \gamma_m\boldsymbol{\beta}_m$. Multiplying both sides by A gives $A\mathbf{R} = r_{zm}\mathbf{1} + \gamma_m A\boldsymbol{\beta}_m$.If $A\boldsymbol{\beta}_m = \boldsymbol{\beta}_m$ then $A\mathbf{R}$ would have to be equal to $\mathbf{R}$, which implies no repackaging ($A = I$).

[16] When $\mathbf{p}$ is efficient the $\mathbf{e}_p$ will be perfectly correlated with $\boldsymbol{\beta}_p$.

$$\gamma_m \begin{pmatrix} \bar{\beta}_m - \bar{\beta}_p \frac{\hat{cov}(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{\hat{var}(\boldsymbol{\beta}_p)} \\ \frac{\hat{cov}(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{\hat{var}(\boldsymbol{\beta}_p)} - 1 \end{pmatrix} \equiv \gamma_m \begin{pmatrix} a \\ b \end{pmatrix}$$

The OLS R^2 is less than one as long as $\mathbf{e}_p - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}_p$ is non-zero. We have

$$\begin{aligned} \mathbf{e}_p - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}_p &= \mathbf{e}_p - \gamma_m \begin{pmatrix} \mathbf{1} & \boldsymbol{\beta}_p \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \mathbf{e}_p - \gamma_m \left(a\mathbf{1} + b\boldsymbol{\beta}_p\right) \\ &= \gamma_m(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p) - \gamma_m \left(a\mathbf{1} + b\boldsymbol{\beta}_p\right) \\ &= \gamma_m \left[\boldsymbol{\beta}_m - (1+b)\boldsymbol{\beta}_p - a\mathbf{1}\right] \neq 0_N \end{aligned}$$

The result holds because $\boldsymbol{\beta}_m$ and $\boldsymbol{\beta}_p$ are not perfectly correlated. In deriving the above result I used the fact that $\mathbf{e}_p = \gamma_m(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p)$, which obtains because $\mathbf{m}$ and $\mathbf{p}$ have the same variance. The arbitrariness of the R^2_{ols} follows from the above repackaging argument.

Appendix 2. The OLS bias.

The OLS estimator is given by

$$\boldsymbol{\Gamma}_{OLS} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{R} = \boldsymbol{\Gamma} + (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}_p$$

We have

$$(\mathbf{X}'\mathbf{X})^{-1} = \frac{1}{N\boldsymbol{\beta}_p'\boldsymbol{\beta}_p - (\mathbf{1}'\boldsymbol{\beta}_p)^2} \begin{pmatrix} \boldsymbol{\beta}_p'\boldsymbol{\beta}_p & -\mathbf{1}'\boldsymbol{\beta}_p \\ -\mathbf{1}'\boldsymbol{\beta}_p & N \end{pmatrix}$$

and

$$\mathbf{X}'\mathbf{e}_p = \gamma_m \begin{pmatrix} \mathbf{1}'(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p) \\ \boldsymbol{\beta}_p'(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p) \end{pmatrix}$$

The bias is given by

$$
\begin{aligned}
(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}_p &= \frac{\gamma_m}{N\boldsymbol{\beta}_p'\boldsymbol{\beta}_p - (\mathbf{1}'\boldsymbol{\beta}_p)^2} \begin{pmatrix} \boldsymbol{\beta}_p'\boldsymbol{\beta}_p & -\mathbf{1}'\boldsymbol{\beta}_p \\ -\mathbf{1}'\boldsymbol{\beta}_p & N \end{pmatrix} \begin{pmatrix} \mathbf{1}'(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p) \\ \boldsymbol{\beta}_p'(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p) \end{pmatrix} \\
&= \frac{\gamma_m}{N\boldsymbol{\beta}_p'\boldsymbol{\beta}_p - (\mathbf{1}'\boldsymbol{\beta}_p)^2} \begin{pmatrix} \mathbf{1}'\boldsymbol{\beta}_m\ \boldsymbol{\beta}_p'\boldsymbol{\beta}_p\ - \mathbf{1}'\boldsymbol{\beta}_p\ \boldsymbol{\beta}_p'\boldsymbol{\beta}_m \\ -\mathbf{1}'\boldsymbol{\beta}_p(\mathbf{1}'\boldsymbol{\beta}_m - \mathbf{1}'\boldsymbol{\beta}_p) + N \times (\boldsymbol{\beta}_p'\boldsymbol{\beta}_m - \boldsymbol{\beta}_p'\boldsymbol{\beta}_p) \end{pmatrix} \\
&= \frac{\gamma_m}{v\hat{a}r(\boldsymbol{\beta}_p)} \begin{pmatrix} \bar{\beta}_m v\hat{a}r(\boldsymbol{\beta}_p) - \bar{\beta}_p c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m) \\ c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m) - v\hat{a}r(\boldsymbol{\beta}_p) \end{pmatrix} \\
&= \gamma_m \begin{pmatrix} \bar{\beta}_m - \bar{\beta}_p \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} \\ \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} - 1 \end{pmatrix}
\end{aligned}
$$

When $\mathbf{p}$ is inefficient and $\sigma_p^2 \neq \sigma_m^2$, we have the same model except $\boldsymbol{\Gamma} = [r_{zm} \quad \gamma_p]'$, $\gamma_p = \gamma_m \sigma_p^2/\sigma_m^2$, and $\mathbf{e}_p = \gamma_m(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p \sigma_p^2/\sigma_m^2) = \gamma_m(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p^*)$. While the matrix $(\mathbf{X}'\mathbf{X})^{-1}$ is unchanged, the vector $\mathbf{X}'\mathbf{e}_p$ is now given by

$$
\mathbf{X}'\mathbf{e}_p = \gamma_m \begin{pmatrix} \mathbf{1}'(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p^*) \\ \boldsymbol{\beta}_p'(\boldsymbol{\beta}_m - \boldsymbol{\beta}_p^*) \end{pmatrix}
$$

Repeating the same steps gives

$$
(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{e}_p = \gamma_m \begin{pmatrix} \bar{\beta}_m - \bar{\beta}_p \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} \\ \frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} - \sigma_p^2/\sigma_m^2 \end{pmatrix}
$$

The intercept bias is the same, while the slope bias is now

$$
\begin{aligned}
\hat{\gamma}_m(ols) &= \gamma_p + \gamma_m \left(\frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} - \sigma_p^2/\sigma_m^2 \right) \\
&= \gamma_m + \gamma_m \left(\frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} - \sigma_p^2/\sigma_m^2 + (\sigma_p^2/\sigma_m^2 - 1) \right) \\
&= \gamma_m + \gamma_m \left(\frac{c\hat{o}v(\boldsymbol{\beta}_p, \boldsymbol{\beta}_m)}{v\hat{a}r(\boldsymbol{\beta}_p)} - 1 \right)
\end{aligned}
$$

which is identical to the case where $\sigma_p^2 = \sigma_m^2$.

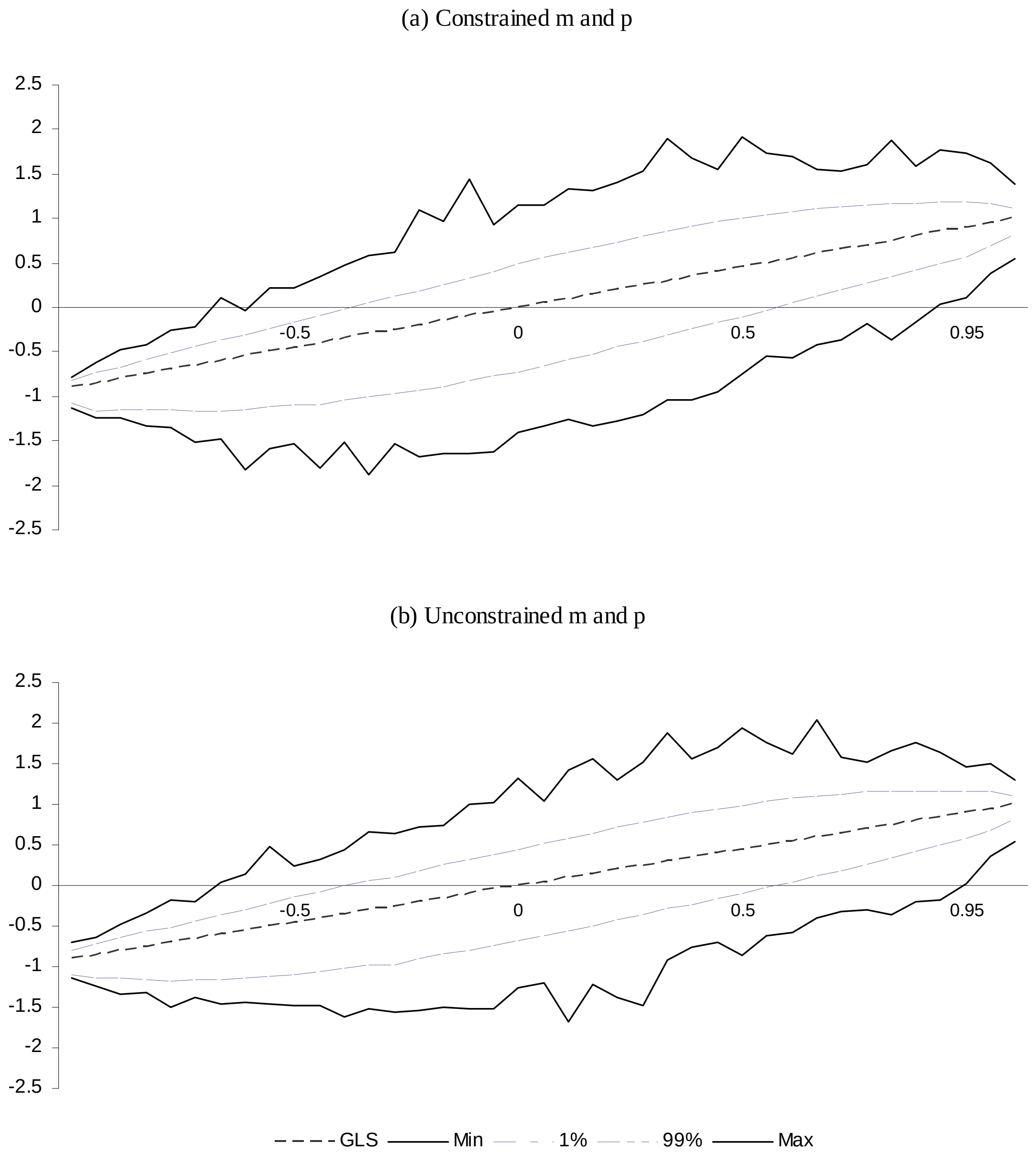
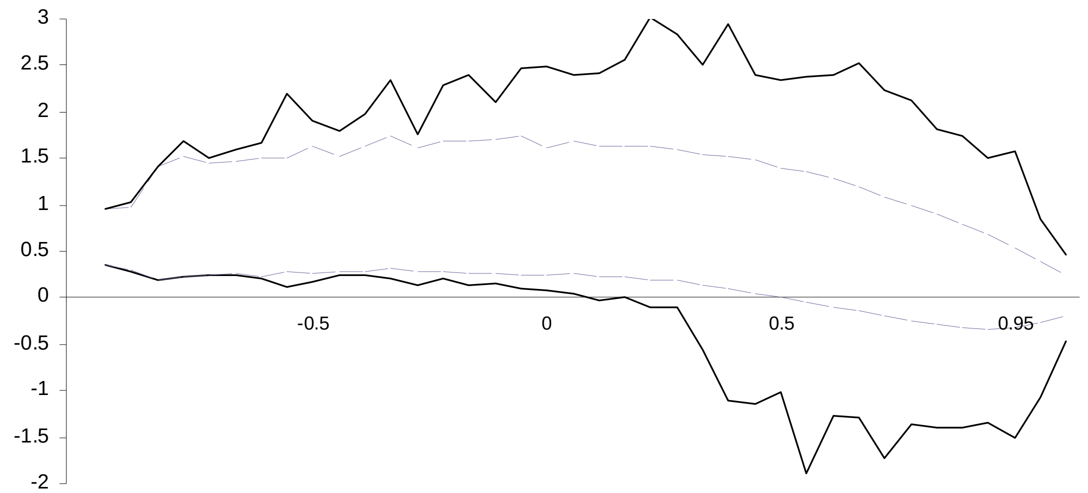


Figure 1. The relative bias of GLS and OLS slope estimates. The data are based on 1000 simulated frontiers, each with 1000 departures from the market portfolio (1 million replications). The plotted percentiles are based on 39 categories of efficiency, for values of ψ_p between -1 and up to, but excluding, 1. The graphs show the minimum, the maximum, and the 1st and 99th percentiles. The straight line is the GLS relative bias. The vertical axis is the relative bias, which is plotted against the relative measure of efficiency, ψ_p. In the constrained case both the market and the inefficient portfolios have positive weights. The unconstrained case allows the market and the inefficient portfolios to have short positions.

(a) Constrained m and p



(b) Unconstrained m and p

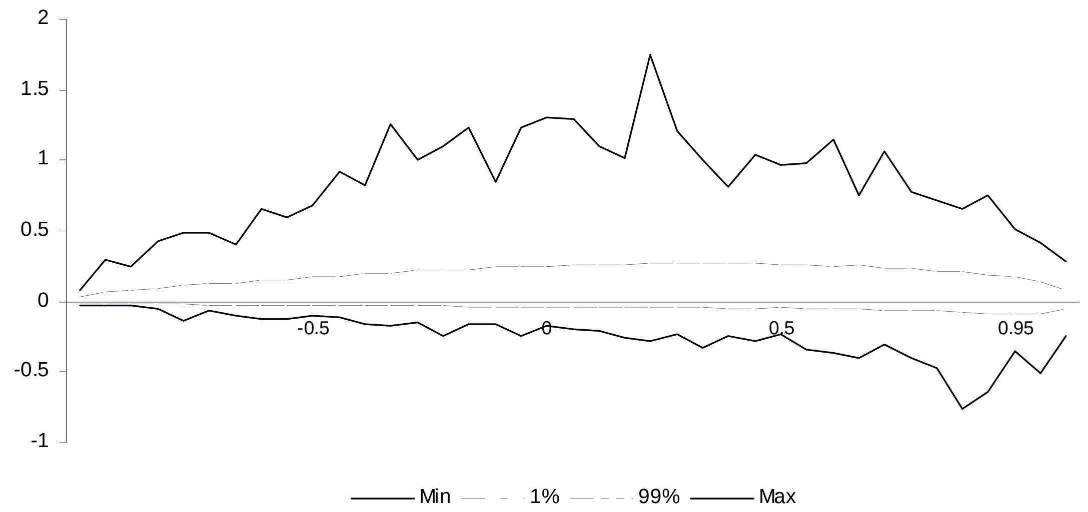


Figure 2. The relative bias of OLS intercept estimates. The data are based on 1000 simulated frontiers, each with 1000 departures from the market portfolio (1 million replications). The plotted percentiles are based on 39 categories of efficiency, for values of ψ_p between -1 and up to, but excluding, 1. The graphs show the minimum, the maximum, and the 1^{st} and 99^{th} percentiles of the OLS bias. The vertical axis is the relative bias, which is plotted against the relative measure of efficiency, ψ_p. In the constrained case both the market and the inefficient portfolios have positive weights. The unconstrained case allows the market and the inefficient portfolios to have short positions.

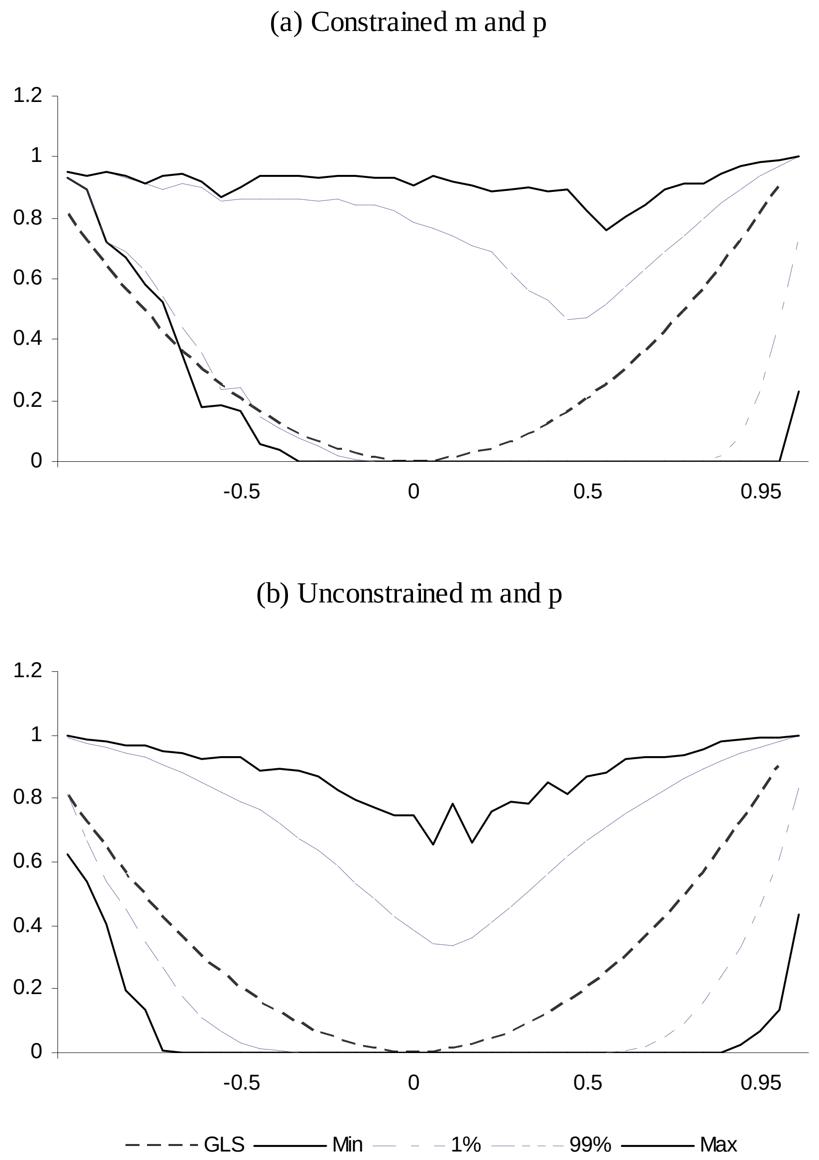


Figure 3. The OLS and GLS R-squared. The data are based on 1000 simulated frontiers, each with 1000 departures from the market portfolio (1 million replications). The plotted percentiles are based on 39 categories of efficiency, for values of ψ_p between -1 and up to, but excluding, 1. The graphs show the minimum, the maximum, and the 1^{st} and 99^{th} percentiles of the OLS R-squared. The middle curve is the GLS R-squared. In the constrained case both the market and the inefficient portfolios have positive weights. The unconstrained case allows the market and the inefficient portfolios to have short positions.

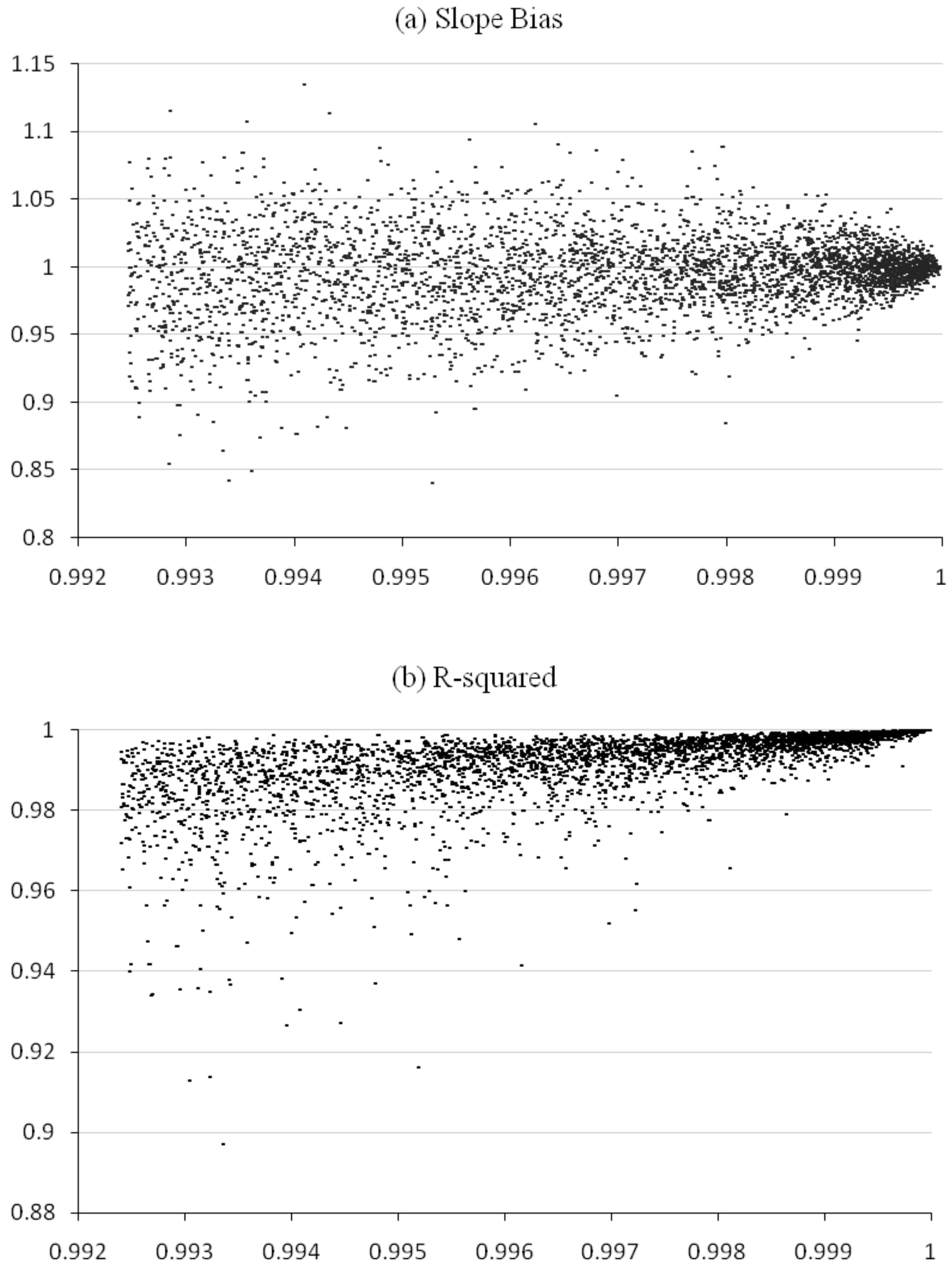
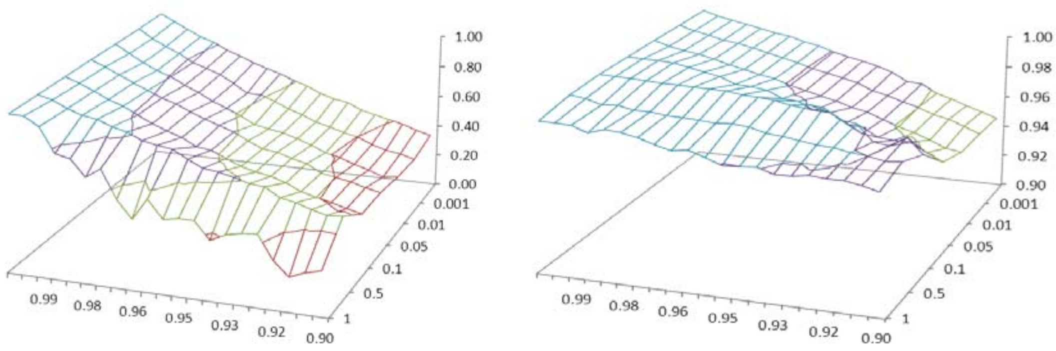
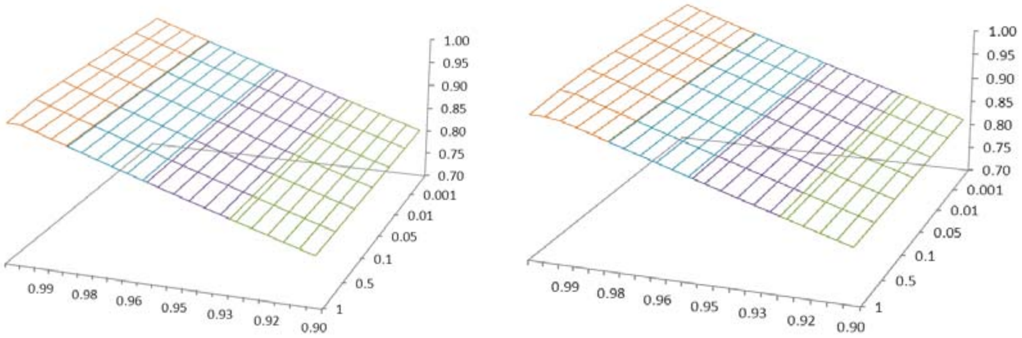


Figure 4. OLS slope bias and OLS R-squared for the top 4000 efficient portfolios (constrained m and p). The data are based on the 1 million replications described in Figures 1 and 3. The horizontal axis shows the values of ψ_p (the efficiency of the index used to compute beta).



OLS (1st Percentile) (99th Percentile)



GLS (1st Percentile) (99th Percentile)

Figure 5. OLS and GLS R-squared for the nearly efficient portfolios (constrained m and p). The vertical axis represents the R-Squared. The omitted variable scale takes the values (0, 0.001, 0.01, 0.05, 0.1, 0.5, 1). For each scale, the most efficient portfolios are obtained from 1 million replications as described in Figures 1 and 3. The values of ψ_p (the efficiency of the index used to compute beta) vary between 0.90 and 1.

Table 1. R-Squared statistics for selected efficiency and omitted sensitivity levels.

Panel A: OLS		Scale of Omitted Sensitivities						
Efficiency Range	Percentile	0	0.001	0.01	0.05	0.1	0.5	1
0.895-0.900	1%	0.331	0.293	0.305	0.336	0.337	0.460	0.277
	99%	0.945	0.947	0.947	0.953	0.971	0.979	0.974
0.945-0.950	1%	0.606	0.577	0.540	0.613	0.620	0.691	0.460
	99%	0.977	0.977	0.977	0.983	0.995	0.991	0.983
0.975-0.980	1%	0.812	0.780	0.795	0.819	0.833	0.870	0.754
	99%	0.991	0.991	0.991	0.997	0.998	0.993	0.993
0.985-0.990	1%	0.885	0.869	0.869	0.899	0.908	0.914	0.924
	99%	0.995	0.995	0.995	0.999	0.999	0.994	0.994
0.990-0.995	1%	0.930	0.914	0.920	0.940	0.947	0.957	0.977
	99%	0.997	0.997	0.997	1.000	0.999	0.997	0.994
0.995-1.00	1%	0.973	0.972	0.973	0.977	0.978	0.979	na
	99%	1.000	1.000	1.000	1.000	1.000	0.998	na

Panel B: GLS		Scale of Omitted Sensitivities						
Efficiency Range	Percentile	0	0.001	0.01	0.05	0.1	0.5	1
0.895-0.900	1%	0.801	0.801	0.801	0.801	0.801	0.801	0.801
	99%	0.810	0.810	0.810	0.810	0.810	0.810	0.810
0.945-0.950	1%	0.893	0.893	0.893	0.893	0.893	0.893	0.893
	99%	0.902	0.902	0.902	0.902	0.902	0.902	0.902
0.975-0.980	1%	0.951	0.951	0.951	0.951	0.951	0.951	0.951
	99%	0.960	0.960	0.960	0.960	0.960	0.960	0.960
0.985-0.990	1%	0.970	0.970	0.970	0.970	0.970	0.970	0.970
	99%	0.980	0.980	0.980	0.980	0.980	0.980	0.979
0.990-0.995	1%	0.980	0.980	0.980	0.980	0.980	0.980	0.981
	99%	0.990	0.990	0.990	0.990	0.990	0.989	0.984
0.995-1.00	1%	0.990	0.990	0.990	0.990	0.990	0.990	na
	99%	1.000	1.000	1.000	0.999	0.999	0.994	na

Note: The percentiles are based on simulated data as described in Figure 5.