On the Explanatory Power of Asset Pricing Models Across and Within Portfolios

RAYMOND KAN*

Comments are welcome

January, 1999

*University of Toronto and currently visiting University of California at Irvine. I thank Social Sciences and Humanities Research Council of Canada for financial support. I am grateful to Jonathan Berk, Alexandar Mackay, Guofu Zhou, and particularly Chu Zhang, for many helpful comments and suggestions. All errors are mine. Address correspondence to Raymond Kan, UCI Graduate School of Management, University of California, Irvine, CA 92697-3125, U.S.A., Phone: (949)824-8365, E-mail: rkan@gsm.uci.edu.

On the Explanatory Power of Asset Pricing Models Across and Within Portfolios

ABSTRACT

In this paper, we investigate the effect of using portfolios in assessing the explanatory power of an asset pricing model, where the portfolios are formed by sorting firms using a firm-specific variable. We show that the explanatory power of an asset pricing model at the individual firm level can be grossly exaggerated or nullified at the portfolio level, depending on the choice of the sorting variable. We also study the explanatory power of an asset pricing model on the firms within each portfolio, and show that in general the explanatory power of an asset pricing model for firms within a portfolio can increase or decrease with the number of portfolios.

Although asset pricing models are supposed to work for individual firms as well as for portfolios, they are often estimated and tested using portfolios. There are various reasons for using portfolios to assess asset pricing models. While most of the reasons are statistical in nature and procedures to address some of them are available, the use of portfolios in tests of asset pricing models is still prevalent. Since for most applications, we are interested in how good an asset pricing model is at explaining the expected returns of individual firms, it is important for us to know what we can learn about asset pricing relations at the individual firm level from asset pricing studies that are based on portfolios. Does the performance of an asset pricing model at the portfolio level tell us anything about its performance at the individual firm level? In recent literature, there are also studies that examine the performance of asset pricing models within the firms of each portfolio. Unless an asset pricing model is perfect, the performance of it within the portfolios will be, in general, different from that in the entire sample. It is of interest for us to understand the theoretical relation between these two measures of performance. Does the performance of an asset pricing model within the portfolios appear to be better or worse than that in the whole sample?

Cautions regarding use of portfolios to test asset pricing models abound in the literature. Roll (1977), for example, suggests that by forming portfolios, pricing errors of individual firms can disappear in portfolios, and hence tests based on portfolios may produce supporting results even when the model is false. Grauer and Janmatt (1988) provide the condition for the pricing errors of individual firms to disappear in portfolios, and also provide an example to show that even when the CAPM has no explanatory power for expected returns of individual firms, it can perfectly explain the expected returns of some portfolios. While using portfolios can make a bad asset pricing model look good, it could also make a good asset pricing model look bad. Lo and MacKinlay (1990) suggest that when portfolios are formed based on a sorting variable that is known to be correlated with *ex post* pricing errors, then the asset pricing model will be over-rejected at the portfolio level even though it is true. Kandel and Stambaugh (1995) show that even though the CAPM is almost true on a set of firms, there exists some (repackaged) portfolios of the firms on which the CAPM has almost no explanatory power.

In this paper, we focus on portfolios that are formed based on sorting of a firm-specific variable. We address the issue of how the theoretical explanatory power of an asset pricing model on such portfolios is related to its explanatory power at the individual firm level as well as the sorting variable. On an *ex ante* basis, without knowing the pricing errors of an asset pricing model on individual firms, it is not entirely clear how sorting could systematically improve or destroy the explanatory power of an asset pricing model at the portfolio level. Our analysis, however, shows that the explanatory power of an asset pricing model on such portfolios is only determined by the sorting variable. The explanatory power of an asset pricing model at the individual firm level plays absolutely no role in determining the explanatory power of the model at the portfolio level. In the extreme cases, an asset pricing model at the portfolio level can be either perfect or totally incapable of explaining the expected returns, regardless of how good or how bad the asset pricing model is at the individual firm level. These results cast serious doubts on what we really can learn from empirical asset pricing studies that use portfolios.

In addition, we address the issue of how the explanatory power of an asset pricing model is affected if we examine it only using firms within a portfolio. In a recent paper, Berk (1998) suggests that increasing the number of portfolios would decrease the explanatory power of an asset pricing model within each portfolio. Our analysis suggests that his result does not hold in general. In the case where the sorting variable is insufficient to fully describe expected return, increasing the number of portfolios can increase or decrease the explanatory power of an asset pricing model on the firms within a portfolio.

The rest of the paper is organized as follows. Section I discusses the explanatory power of an asset pricing model for individual firms, across portfolios, and within portfolios. It provides an analysis of the relation between the explanatory power of an asset pricing model at the individual firm level and at the portfolio level. It also provides an analysis of how the explanatory power of an asset pricing model for individual firms within a portfolio changes with the number of portfolios. Section II provides simulation evidence to illustrate our theoretical results in some realistic settings. The final section provides our conclusions and the Appendix contains proofs of all propositions.

I. The Theory

A. Explanatory Power of an Asset Pricing Model for Individual Firms

We assume there are N individual firms in our sample, where these firms are considered to be independently drawn from a much larger population, so that the N firms that we choose are representative of the firms in the population. Each firm in the population is characterized by a triplet $y = [\mu, m, s]'$ where μ is the true expected return of the firm, m is the predicted expected return of the firm based on an asset pricing model,[1] and s is the value of a firm-specific variable that is used to sort firms into portfolios. We assume the cross-sectional distribution of this triplet in the population is continuous with density function $f_y(y)$. We also assume the mean and the variance-covariance matrix of this distribution exist, and the variance-covariance is denoted as

$$\Sigma_y = E[(y - E[y])(y - E[y])'] \equiv \begin{bmatrix} \sigma_\mu^2 & \sigma_{\mu m} & \sigma_{\mu s} \\ \sigma_{\mu m} & \sigma_m^2 & \sigma_{ms} \\ \sigma_{\mu s} & \sigma_{ms} & \sigma_s^2 \end{bmatrix}. \quad (1)$$

Note that Σ_y is assumed to be positive semidefinite but it does not have to be nonsingular. For example, if the asset pricing model is perfect, i.e., $\mu = m$, or the firm-specific variable s completely explains the expected return μ, or the sorting variable is a linear transformation of the predicted expected return, then Σ_y is singular.

[1] For our purpose, an asset pricing model is a model that generates a prediction of expected returns for individual firms. It does not just include models that are motivated by theory. It also includes those that are empirically motivated. Furthermore, m does not have to be the predicted expected return, it could be just a linear transformation of the predicted expected return. For example, we can use β as m for the case of the CAPM.

Since each firm is viewed as a random drawing from the population, it is natural to define the theoretical explanatory power of the asset pricing model on the expected returns of individual firms as the squared correlation coefficient between the true expected return μ and the predicted expected return from the model m:

$$\rho^2_{\mu m} = \frac{\sigma^2_{\mu m}}{\sigma^2_\mu \sigma^2_m}. \tag{2}$$

Note that this theoretical measure is computed based on the population moments. It is different from the measure of explanatory power that is computed using just the N individual firms in the sample, which we define as

$$R^2_{\mu m} = \frac{\left[\sum_{i=1}^N (\mu_i - \bar{\mu})(m_i - \bar{m})\right]^2}{\sum_{i=1}^N (\mu_i - \bar{\mu})^2 \sum_{i=1}^N (m_i - \bar{m})^2}, \tag{3}$$

where $\bar{\mu} = \sum_{i=1}^N \mu_i / N$ and $\bar{m} = \sum_{i=1}^N m_i / N$. This sample measure of explanatory power is similar to the R^2_{OLS} measure defined by Kandel and Stambaugh (1995). However, when $N \to \infty$, $R^2_{\mu m} \to \rho^2_{\mu m}$. In this paper, we are interested in the population measure because we care about not just how good an asset pricing model is at explaining the expected returns of the N firms in the sample, but also how good it is at explaining the expected returns of all the other firms that are not included in the sample.

Before we move on, we would like to clarify what are random variables and what are constants in our model. In our context, the triplet y is considered a random variable before we select the N firms from the population. But once the N firms in our sample are chosen, $y_1, y_2, \ldots, y_N$ are treated as constants. However, of the three elements of y, the first element (μ) is typically not observable by the econometrician and hence the measures $\rho^2_{\mu m}$ and $R^2_{\mu m}$ are only theoretical and are not directly observable. In practice, one can use the average return ($\bar{R}$) as a proxy for the expected return and compute $R^2_{\bar{R}m}$, using $\bar{R}$ in lieu of μ. However, since the value of $R^2_{\bar{R}m}$ depends on the realizations of average returns, it is a random variable even after we have chosen the N firms in our sample.[2] As the number of time series observations used in computing $\bar{R}$

[2]The fact firms are random samples from the same population does not preclude correlation between

increases, we have $\bar{R} \to \mu$ and hence the measure of explanatory power that is computed using average returns will converge to the theoretical measure that is computed based on expected returns. In order to better focus on our main issue, we will not consider measures that are based on average returns in this paper.[3]

It is important to realize that while our definition of theoretical explanatory power is a reasonable one and used by many others, it is not a perfect one. There exist also other measures of the theoretical explanatory power (see Chen, Kan, and Zhang (1998) for a discussion of various measures) of an asset pricing model but to avoid possible confusion, we will limit our discussion to this particular choice.

B. Explanatory Power of an Asset Pricing Model Across Portfolios

Portfolios are often used in tests of asset pricing models. In this paper we limit our attention to the so-called "equal number, equally weighted" portfolios defined in Kan and Zhang (1995). Namely, the N firms are sorted by s in ascending order into n portfolios, where $n \geq 2$. The n portfolios are nonoverlapping (i.e., each firm can belong to only one portfolio) and each contains (roughly) the same number of firms. Within each portfolio, the returns as well as the firm-specific variables are equally weighted. Under this scheme of portfolio formation, the population of the ith portfolio consists of all the firms with $s^*_{i-1} \leq s < s^*_i$, where $\int_{-\infty}^{s^*_i} f_s(s) = \frac{i}{n}$.[4] Conditional on a firm belongs to portfolio i, the joint distribution of its triplet y is given by

$$f_y(y|s^*_{i-1} \leq s < s^*_i) = \begin{cases} nf_y(y) & \text{if } s^*_{i-1} \leq s < s^*_i, \\ 0 & \text{otherwise.} \end{cases} \quad (4)$$

realized returns on two firms. Realized returns on all firms are their expected returns plus unexpected returns; the latter are random variables defined on another probability space.

[3]Conditional on N firms being chosen, Chen, Kan, and Zhang (1998) provide an analysis of the sampling distribution of $R^2_{\bar{R}m}$.

[4]In practice, the cutoff points of the n portfolios are determined by the firm-specific variables of the N firms in the sample. We assume N is large enough that the population cutoff points are good approximations of the sample cutoff points.

Since the portfolio is equally weighted, the theoretical values of the triplet y for portfolio i are simply their conditional means. Therefore, we have

$$\begin{aligned}
\mu_p^i &= E[\mu|s_{i-1}^* \leq s < s_i^*] \\
&= n\int_{s_{i-1}^*}^{s_i^*}\int_{-\infty}^{\infty}\mu f_{\mu,s}(\mu,s)\mathrm{d}\mu\mathrm{d}s \\
&= n\int_{s_{i-1}^*}^{s_i^*}E[\mu|s]f_s(s)\mathrm{d}s, \qquad (5)\\
m_p^i &= E[m|s_{i-1}^* \leq s < s_i^*] \\
&= n\int_{s_{i-1}^*}^{s_i^*}\int_{-\infty}^{\infty}m f_{m,s}(m,s)\mathrm{d}m\mathrm{d}s \\
&= n\int_{s_{i-1}^*}^{s_i^*}E[m|s]f_s(s)\mathrm{d}s, \qquad (6)\\
s_p^i &= E[s|s_{i-1}^* \leq s < s_i^*] \\
&= n\int_{s_{i-1}^*}^{s_i^*}s f_s(s)\mathrm{d}s. \qquad (7)
\end{aligned}$$

Similar to the measure of explanatory power for individual firms, we define the theoretical explanatory power of the asset pricing model on the expected returns of these n portfolios as the squared correlation between μ_p^i and m_p^i:

$$\rho_{\mu m}^2(n) = \frac{\left[\sum_{i=1}^n(\mu_p^i - \bar{\mu}_p)(m_p^i - \bar{m}_p)\right]^2}{\sum_{i=1}^n(\mu_p^i - \bar{\mu}_p)^2\sum_{i=1}^n(m_p^i - \bar{m}_p)^2}, \qquad (8)$$

where $\bar{\mu}_p = \sum_{i=1}^n \mu_p^i/n = E[\mu]$ and $\bar{m}_p = \sum_{i=1}^n m_p^i/n = E[m]$.[5]

In comparing expressions (2) and (8), we can determine whether the explanatory power of an asset pricing model is higher or lower for individual firms than for the portfolios. From (5)–(6), we can see that the theoretical explanatory power of an asset pricing model on the n portfolios only depends on (a) the cross-sectional distribution of the sorting variable, $f_s(s)$, (b) the conditional mean of the expected return as a function of the sorting variable, $E[\mu|s]$, and (c) the conditional mean of the predicted expected return by the asset pricing model as a function of the sorting variable, $E[m|s]$. In

[5]Implicitly, we assume the m_p^i are not constant across portfolios. If the μ_p^i are constant across portfolios, $\rho_{\mu m}^2(n)$ is equal to zero by convention.

particular, it does not depend on the joint distribution $f_{\mu,m}(\mu, m)$ and hence $\rho^2_{\mu m}$ at the individual firm level does not play any role in determining $\rho^2_{\mu m}(n)$ at the portfolio level.[6] The only possible exception is when we sort the portfolios using $s = m$. However, even when $\rho^2_{sm} = 1$, there is still no obvious relation between $\rho^2_{\mu m}$ and $\rho^2_{\mu m}(n)$ since the former depends on the entire joint distribution $f_{\mu,m}(\mu, m)$ whereas the latter only depends on the conditional expectation $E[\mu|m]$ and the marginal distribution $f_m(m)$. Our observation that $\rho^2_{\mu m}(n)$ at the portfolio level has no relation to $\rho^2_{\mu m}$ at the individual firm level applies to more general portfolio formation schemes. For example, s could be a vector of firm-specific variables and portfolios could be sorted on a multi-dimensional basis using different combinations of elements of s. Moreover, the portfolios do not have to be equally weighted, and they do not need to have equal number of firms.[7]

The observation that $\rho^2_{\mu m}(n)$ has nothing to do with $\rho^2_{\mu m}$ is surprising. It suggests that we cannot infer the explanatory power of an asset pricing model at the individual firm level (which is presumably about what we care) from empirical studies that only use portfolio data. The explanatory power of an asset pricing model on the portfolios only depends on the sorting variable and its relation with μ and m. Different sorting schemes could possibly provide a wide range of $\rho^2_{\mu m}(n)$, but none of them can be relied on to provide information about $\rho^2_{\mu m}$. In the following, we focus on the two extreme situations in which the explanatory power of an asset pricing model at the portfolio level can appear very high or very low.

Proposition 1 *If there exist scalars a and b such that $E[\mu|s] = a + bE[m|s]$, then for*

[6] In general, we can write $\mu = E[\mu|s] + e_1$ and $m = E[m|s] + e_2$ where e_1 and e_2 are conditionally independent of s. $\rho^2_{\mu m}(n)$ depends on only $E[\mu|s]$ and $E[m|s]$ but $\rho^2_{\mu m}$ depends on also e_1 and e_2. For any given value of $\rho^2_{\mu m}(n)$, we can set the covariance matrix of e_1 and e_2 to make $\rho_{\mu m}$ equal to any value in $(-1, 1)$.

[7] If the population has a finite number of firms, then when the number of portfolios is equal to the number of firms, $\rho^2_{\mu m}(n) = \rho^2_{\mu m}$ trivially. In our setup, the population from which our firms come from has a continuous distribution, so this scenario will not occur even for countably infinite number of portfolios.

all $n > 1$,

$$\rho^2_{\mu m}(n) = \begin{cases} 1 & \text{if } b \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{9}$$

An important case is when a sorting variable s is chosen such that the pricing error $\mu - m$ at the individual firm level is conditionally independent of s, i.e., $E[\mu - m|s] = 0$. In that case $E[\mu|s] = E[m|s]$; the condition of Proposition 1 is satisfied with $a = 0$ and $b = 1$. Since pricing errors of individual firms are uncorrelated with the sorting variable, they are averaged out in portfolios, and hence the model becomes perfect at the portfolio level.

Another important case is when the conditional expectations $E(\mu|s)$ and $E(m|s)$ are linear in s. In that case, the conditional expectations, $E(\mu|s)$ and $E(m|s)$, must be linear to each other. We immediately obtain a corollary.

Corollary 1 *Suppose* $E[\mu|s]$ *and* $E[m|s]$ *are linear in* s*. We have for all* $n > 1$,

$$\rho^2_{\mu m}(n) = \begin{cases} 1 & \text{if } \sigma_{\mu s} \neq 0 \text{ and } \sigma_{ms} \neq 0, \\ 0 & \text{otherwise.} \end{cases} \tag{10}$$

This case is also important because it depends merely on the distribution types without reference to the pricing errors. For example, if the true expected return, the predicted expected return, and the firm-specific variable have a multivariate elliptical distribution, then $E[\mu|s]$ and $E[m|s]$ are linear in s.[8]

Proposition 1 and its corollary suggest that when $E[\mu|s]$ and $E[m|s]$ are linear function of each other, the theoretical explanatory power of an asset pricing model at the portfolio level can take only two possible values. It can only be zero or one. The case for $\rho^2_{\mu m}(n) = 0$ is easy to understand. It happens when the sorting variable s is totally uncorrelated with expected returns (or predicted expected returns) of individual firms.

[8]Note that Corollary 1 only requires $E[\mu|s]$ and $E[m|s]$ to be linear in s. It does not require $E[\mu|m]$ to be linear in m (or vice versa), so having multivariate elliptical distribution is only a sufficient but not a necessary condition for Corollary 1 to hold. Furthermore, the condition that $E[\mu|s]$ and $E[m|s]$ are linear in s can be relaxed to the condition that there exists a monotonic function $u = \psi(s)$ such that $E[\mu|u]$ and $E[m|u]$ are linear in u. This is so because sorting by s is the same as sorting by u.

In that case, the resulting portfolios will not display any cross-sectional differences in expected returns (or predicted expected returns). Therefore, it does not matter how good the asset pricing model is, we would not see any explanatory power at the portfolio level.

In practice, it is most likely that one would sort firms into portfolios using a firm-specific variable that would have at least some correlation with the true expected return and possibly some correlation with the predicted expected return as well. In that case, Proposition 1 suggests that as long as $E[\mu|s]$ and $E[m|s]$ are linear functions of each other, the asset pricing model will always perform perfectly at the portfolio level, regardless of the number of portfolios used, and regardless of how poor the asset pricing model actually performs at the individual firm level.[9] This will occur even when the asset pricing model alone has no explanatory power on the expected returns of individual firms, i.e., $\rho^2_{\mu m} = 0$.

The results in Proposition 1 and its corollary are disturbing. It means that finding an asset pricing model that performs perfectly at the portfolio level says nothing about the goodness of that asset pricing model at the individual firm level. That there exists some portfolios on which a wrong asset pricing model can look good is not an entirely new result. However, the results of our Proposition 1 and its corollary are more dramatic than what we can anticipate from other studies. They suggest that one does not need to know the pricing errors of an asset pricing model in order to find portfolios on which the asset pricing model will have perfect explanatory power. All it takes is for us to sort firms using a firm-specific variable s such that $E[\mu|s]$ and $E[m|s]$ are linear functions of each other. Then the asset pricing model will perform perfectly on any number of portfolios that are formed based on such a firm-specific variable.

Proposition 1 also provides clues regarding what situations that an asset pricing model would have poor explanatory power on a set of portfolios. It has to do with the

[9] $\rho^2_{\mu m}(n) = 1$ does not imply $\mu^i_p = m^i_p$. However, to the extent that we can find a prediction that is perfectly correlated with expected returns, it is not difficult to do a linear transformation on it to make $\mu^i_p = m^i_p$.

nature of the sorting variable s. If $E[\mu|s]$ and $E[m|s]$ are not highly correlated with each other, we would find portfolios that are formed based on such a sorting variable s to have low $\rho^2_{\mu m}(n)$. However, the poor performance of the asset pricing model on these portfolios is only due to the choice of the sorting variable, and it says nothing about how good or how bad the asset pricing model actually is at the individual firm level.

Empirically, even when the conditions for Proposition 1 are met, one will still find that an asset pricing model does not provide perfect explanatory power on the average returns of the portfolios. This could be due to the fact that average returns are noisy measures of expected returns. It could also be because the total number of firms in the sample, N, or the number of firms in each portfolio, N/n, is not large enough to give us a good approximation of the population. In Section II, we provide simulation evidence to suggest that the second reason is probably not very important in most empirical studies.

C. Explanatory Power of an Asset Pricing Model for Individual Firms within a Portfolio

There are cases in which researchers would first sort the firms in the sample into portfolios and then examine the validity of the asset pricing model using firms within each one of the portfolios. Berk (1998) provides such an analysis and suggests this procedure is biased in favor of rejecting the asset pricing model under consideration if the sorting variable is correlated with expected return. In this subsection, we refine his analysis and provide conditions under which his claim is true.

If we limit our attention to firms in portfolio i, then the explanatory power of an asset pricing model for firms in this portfolio will be given by

$$\rho^2_{\mu m \cdot i} = \frac{\sigma^2_{\mu m \cdot i}}{\sigma^2_{\mu \cdot i}\sigma^2_{m \cdot i}}, \tag{11}$$

where

$$\sigma^2_{\mu \cdot i} \quad = \quad E[(\mu - \mu^i_p)^2 | s^*_{i-1} \leq s < s^*_i], \tag{12}$$

$$\sigma^2_{m\cdot i} = E[(m - m^i_p)^2 | s^*_{i-1} \leq s < s^*_i], \tag{13}$$

$$\sigma_{\mu m\cdot i} = E[(\mu - \mu^i_p)(m - m^i_p) | s^*_{i-1} \leq s < s^*_i]. \tag{14}$$

Since conditional on $s^*_{i-1} \leq s < s^*_i$, the joint density of μ and m is given by

$$f_{\mu m\cdot i}(\mu, m) = n \int_{s^*_{i-1}}^{s^*_i} f_y(\mu, m, s) \mathrm{d}s, \tag{15}$$

we can rewrite the expressions above as

$$\sigma^2_{\mu\cdot i} = n \int_{s^*_{i-1}}^{s^*_i} \int_{-\infty}^{\infty} (\mu - \mu^i_p)^2 f_{\mu,s}(\mu, s) \mathrm{d}\mu \mathrm{d}s, \tag{16}$$

$$\sigma^2_{m\cdot i} = n \int_{s^*_{i-1}}^{s^*_i} \int_{-\infty}^{\infty} (m - m^i_p)^2 f_{m,s}(m, s) \mathrm{d}m \mathrm{d}s, \tag{17}$$

$$\sigma_{\mu m\cdot i} = n \int_{s^*_{i-1}}^{s^*_i} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mu - \mu^i_p)(m - m^i_p) f_y(\mu, m, s) \mathrm{d}\mu \mathrm{d}m \mathrm{d}s. \tag{18}$$

In Berk (1998), he suggests that as long as s can generate cross-sectional differences in expected returns of the portfolios, then by sorting firms into portfolios, the cross-sectional variations of expected returns within the portfolios will be reduced and hence he claims[10]

$$\rho^2_{\mu m\cdot i} < \rho^2_{\mu m}, \quad i = 1, 2, \ldots, n. \tag{19}$$

For the case when n is large enough, he further claims that

$$\rho^2_{\mu m\cdot i} < \epsilon, \qquad i = 1, 2, \ldots, n, \tag{20}$$

for any arbitrary small $\epsilon > 0$. Based on these claims, he suggests that the explanatory power of an asset pricing model within the firms of a portfolio is a decreasing function of the number of portfolios, and that $\rho^2_{\mu m\cdot i}$ understates the true explanatory power of an asset pricing model.

It is obvious that both claims cannot be true in general. For example, if $\mu = m + s$ where m and s are independent, then $\rho^2_{\mu m} < 1$ at the individual firm level. However,

[10]Berk's (1998) definition of explanatory power of an asset pricing model differs slightly from ours in that it is the cross-sectional R^2 between average returns and the predicted expected returns from the asset pricing model. However, it is clear from his analysis that the use of expected returns instead of average returns will not alter his conclusions at all.

if we sort the firms into portfolios based on s, then there will be little variation of s within each portfolio. Consequently, μ and m will be almost perfectly correlated within the firms of each portfolio as the number of portfolios increases, and we have $\rho^2_{\mu m \cdot i} \approx 1$ for every portfolio. This example shows that explanatory power of an asset pricing model within each portfolio does not have to be lower than its explanatory power for the entire population of firms. Even in the extreme case that the sorting variable is perfectly correlated with true expected returns, these two claims are not always true. For example, if $s = \mu$ and it has an exponential distribution, then it is easy to show that for the last portfolio, we have $\sigma^2_{\mu \cdot n} = \sigma^2_\mu$ and its value is independent of n.[11] If the pricing error $\alpha = \mu - m$ is independent of μ, then we have $\lim_{n\to\infty} \rho^2_{\mu m \cdot n} = \rho^2_{\mu m}$, so the explanatory power of the asset pricing model within the last portfolio will not go to zero and it always stays the same regardless of how many portfolios are formed.[12]

The reason why $\rho^2_{\mu m \cdot i}$ can be greater than $\rho^2_{\mu m}$ is simple. From the familiar variance and covariance decomposition formulae,[13] we have

$$\frac{1}{n}\sum_{i=1}^{n} \sigma^2_{\mu \cdot i} = \sigma^2_\mu - \frac{1}{n}\sum_{i=1}^{N}(\mu_p - \bar{\mu}_p)^2 < \sigma^2_\mu, \tag{21}$$

$$\frac{1}{n}\sum_{i=1}^{n} \sigma^2_{m \cdot i} = \sigma^2_m - \frac{1}{n}\sum_{i=1}^{N}(m_p - \bar{m}_p)^2 < \sigma^2_m, \tag{22}$$

$$\frac{1}{n}\sum_{i=1}^{n} \sigma_{\mu m \cdot i} = \sigma_{\mu m} - \frac{1}{n}\sum_{i=1}^{N}(\mu_p - \bar{\mu_p})(m_p - \bar{m}_p). \tag{23}$$

Therefore, sorting will on average reduce the variation of expected returns and predicted expected returns within a portfolio. Since the variance terms $\sigma^2_{\mu \cdot i}$ and $\sigma^2_{m \cdot i}$ are in the

[11]This result follows from the familiar fact that if X has an exponential distribution, we have $P[X > a + b | X > a] = P[X > b]$ and hence the variance of the truncated exponential distribution is the same as the variance of the unconditional distribution.

[12]Under the same conditions but μ has a t-distribution, we have $\lim_{n\to\infty} \rho^2_{\mu m \cdot 1} = \lim_{n\to\infty} \rho^2_{\mu m \cdot n} = 1$ and the asset pricing model will have almost perfect explanatory power within the first and the last portfolio when n is large enough.

[13]The decomposition formulae state:

$$\begin{aligned} E[\mathrm{Var}[X|Z]] &= \mathrm{Var}[X] - \mathrm{Var}[E[X|Z]], \\ E[\mathrm{Cov}[X,Y|Z]] &= \mathrm{Cov}[X,Y] - \mathrm{Cov}[E[X|Z], E[Y|Z]], \end{aligned}$$

for random variables X, Y, and Z.

denominator of $\rho^2_{\mu m\cdot i}$, their reduction will actually help to increase $\rho^2_{\mu m\cdot i}$. The key question is whether $\sigma^2_{\mu m\cdot i}$ will reduce more to make $\rho^2_{\mu m\cdot i}$ goes down with sorting. From (23), we can see that it depends on the cross-sectional covariance between μ_p and m_p. Even though we assume this term has the same sign as $\sigma_{\mu m}$ and they are positive, it only guarantees the average covariance between μ and m within the portfolios to be less than $\sigma_{\mu m}$. It does not guarantee the average $\sigma^2_{\mu m\cdot i}$ to be less than $\sigma^2_{\mu m}$. This is because we could have $\sigma_{\mu m\cdot i}$ to be all big but have different signs for different portfolios. Therefore, unlike average $\sigma^2_{\mu\cdot i}$ and $\sigma^2_{m\cdot i}$, the average $\sigma^2_{\mu m\cdot i}$ does not have to decrease with sorting.

Knowing that Berk's claims are not true in general, we proceed to characterize the conditions for which his claims are true. However, analytical expressions for the general case are difficult to derive. Moreover, for a general distribution of y, its mean and variance-covariance matrix are not sufficient to characterize the entire distribution. Therefore, we focus on just the case that y has a multivariate elliptical distribution in the following proposition.[14] By sacrificing generality, we are able to provide analytical results as well as better intuition.

Proposition 2 *If y has a multivariate elliptical distribution with finite mean and variance, we have*

$$\rho^2_{\mu m\cdot i} = \frac{\left[\rho_{\mu m} - \rho_{\mu s}\rho_{ms}h(i)\right]^2}{\left[1-\rho^2_{\mu s}h(i)\right]\left[1-\rho^2_{ms}h(i)\right]}, \tag{24}$$

where

$$h(i) = 1 - \frac{\int_{s^*_{i-1}}^{s^*_i} n\left(\frac{s-s^i_p}{\sigma_s}\right)^2 f_s(s)\mathrm{d}s}{\int_{s^*_{i-1}}^{s^*_i} ng(s)f_s(s)\mathrm{d}s} \tag{25}$$

and $g(s)$ is a positive function which depends on the class of elliptical distribution. For

[14]y is said to have an elliptical distribution if

$$f_y(y) \propto |\Sigma_y|^{-\frac{1}{2}}\varphi((y-E[y])'\Sigma_y^{-1}(y-E[y])),$$

where $\varphi(u)$ is a positive and nonincreasing function of u for $u \geq 0$. See Kelker (1970) for a discussion of various properties of multivariate elliptical distribution. For Proposition 2, all we need is there exists a monotonic function $u = \psi(s)$ such that (μ, m, u) has a multivariate elliptical distribution.

the special case that y has a multivariate normal distribution, we have $g(s) = 1$ and

$$h(i) = n^2 \left[\phi(c_{i-1}^*) - \phi(c_i^*)\right]^2 - n \left[c_{i-1}^* \phi(c_{i-1}^*) - c_i^* \phi(c_i^*)\right], \tag{26}$$

where $c_i^* = (s_i^* - E[s])/\sigma_s$ and $\phi(\cdot)$ is the density function of the standard normal distribution.

Proposition 2 suggests that the explanatory power of an asset pricing model within a portfolio depends on $h(i)$, which is determined by the distribution of s as well as the number of portfolios. In addition, it also depends on $\rho_{\mu m}$, $\rho_{\mu s}$, and ρ_{ms}. Therefore, it is not just the explanatory power of the asset pricing model at the individual firm level that matters. The sorting variable also matters in determining the performance of an asset pricing model within the firms of a portfolio. We first consider a couple of special cases:

1. $\rho_{ms} = 0$. In this case, we have

$$\rho_{\mu m \cdot i}^2 = \frac{\rho_{\mu m}^2}{\left[1 - \rho_{\mu s}^2 h(i)\right]}.$$

   For the case of elliptical distribution, it can be shown that $0 < h(i) < 1$.[15] Therefore, as long as $\rho_{\mu s} \neq 0$, we have $\rho_{\mu m \cdot i}^2 > \rho_{\mu m}^2$ and the explanatory power of an asset pricing model within each portfolio is always greater than its explanatory power on all the firms in the population. While sorting may reduce the cross-sectional variations of expected returns within a portfolio, it does not reduce the covariance between μ and m enough for $\rho_{\mu m \cdot i}$ to go down. Similarly, if $\rho_{\mu s} = 0$, we have $\rho_{\mu m \cdot i}^2 \geq \rho_{\mu m}^2$.

2. $\rho_{\mu m} = 0$. In this case, as long as $\rho_{\mu s} \neq 0$ and $\rho_{ms} \neq 0$, we have

$$\rho_{\mu m \cdot i}^2 = \frac{\rho_{\mu s}^2 \rho_{ms}^2 h(i)^2}{\left[1 - \rho_{\mu s}^2 h(i)\right]\left[1 - \rho_{ms}^2 h(i)\right]} > 0$$

[15]Proof is available upon request.

and the asset pricing model will always have some explanatory power within the firms in each portfolio, even though the asset pricing model alone is completely incapable of explaining expected returns of the firms in the population. This is another case demonstrating that sorting can actually make a poor asset pricing model look good within the portfolios.

3. $\rho_{\mu s} = \pm 1$. In this case, we have $\rho^2_{\mu m} = \rho^2_{ms}$ and

$$\rho^2_{\mu m \cdot i} = \frac{\rho^2_{\mu m}\left[1 - h(i)\right]}{\left[1 - \rho^2_{\mu m} h(i)\right]}.$$

   Therefore, we have $\rho^2_{\mu m \cdot i} < \rho^2_{\mu m}$ for all the portfolios if $\rho^2_{\mu m} \neq 1$. This example was used by Berk (1998) to show that as long as an asset pricing model is not perfect, its explanatory power will be lower within portfolios than for all the firms in the population. However, as discussed earlier, this is not always true for a distribution that is outside of the class of multivariate elliptical distribution. Therefore, in general, one cannot conclude that sorting will always reduce the $\rho^2_{\mu m \cdot i}$ for every portfolio even when the sorting variable is perfectly correlated with expected return.

In general, when the number of portfolios is finite, $\rho^2_{\mu m \cdot i}$ varies across portfolios and hence comparison between $\rho^2_{\mu m}$ and $\rho^2_{\mu m \cdot i}$ is somewhat difficult. However, as the number of portfolios n goes to infinity,[16] $\rho^2_{\mu m \cdot i}$ converges to a common limit for most of the portfolios and hence we can easily address the question as to whether sorting will eventually improve or reduce the explanatory power of an asset pricing model. This result is given in the following proposition.

Proposition 3 *Suppose y has a multivariate elliptical distribution with finite mean and variance. For any $\epsilon_1 > 0$ and $\epsilon_2 > 0$, there exists an n_0 such that for all $n > n_0$, we have $h(i) > 1 - \epsilon_1$ for at least $n(1 - \epsilon_2) - 2$ portfolios. For the case that y has a multivariate*

[16]We can let n goes to infinity because the population is characterized by a continuous distribution and it has uncountably infinite number of firms.

normal distribution, we have a stronger result that $h(i) > 1 - \epsilon_1$ *for all portfolios when* $n > n_0$.

Since $h(i) \leq 1$, Proposition 3 suggests that for most of the portfolios, we have $h(i) \to 1$ and

$$\lim_{n\to\infty} \rho^2_{\mu m \cdot i} = \frac{(\rho_{\mu m} - \rho_{\mu s}\rho_{ms})^2}{(1-\rho^2_{\mu s})(1-\rho^2_{ms})} \equiv \rho^2_{\mu m|s}. \tag{27}$$

Proposition 3 does not suggest that all portfolios will have $\rho^2_{\mu m \cdot i}$ approaching to this limit (except for the case of normal), only that a majority of them will. However, as $n \to \infty$, the set of portfolios that does not approach this limit is of measure zero.[17] The limit $\rho^2_{\mu m|s}$ is the familiar measure of partial coefficient of determination. It measures the squared correlation of μ and m, conditioned on a particular value of the sorting variable s. For the case of multivariate elliptical distribution, this partial coefficient of determination is independent of s. As $n \to \infty$, the cutoff point s^*_{i-1} will converge to s^*_i for most of the portfolios. Therefore, it is not surprising to find that for most of the portfolios, the within-portfolio explanatory power of the asset pricing model will converge to this common limit of $\rho^2_{\mu m|s}$.

From the expression of $\rho^2_{\mu m|s}$, we can see that unless the sorting variable is perfectly correlated with expected return or $\rho_{\mu m} = \rho_{\mu s}\rho_{ms}$,[18] sorting will not completely destroy the explanatory power of an asset pricing model within the portfolios even when the number of portfolios goes to infinity. We therefore focus on the question of when increasing the number of portfolios will reduce the explanatory power of an asset pricing model within the portfolios. The following proposition provides the condition for this to happen.

Proposition 4 *Suppose* y *has a multivariate elliptical distribution with finite mean and*

[17] For example, if y has a multivariate t-distribution with ν ($\nu > 2$) degrees of freedom, we have $\lim_{n\to\infty} h(1) = \lim_{n\to\infty} h(n) = \frac{\nu-2}{\nu-1} < 1$. Sufficient conditions for all the portfolios to approach the limit $\rho^2_{\mu m|s}$ are available upon request.

[18] When μ and s are perfectly correlated, then conditional on s, there is no cross-sectional variation in μ and hence $\rho_{\mu m|s} = 0$ by convention.

variance. For fixed $\rho_{\mu s}$ and ρ_{ms}, we have

$$\begin{array}{llll} \rho^2_{\mu m|s} \leq \rho^2_{\mu m} & \text{if} & \rho_{\mu m} & \text{is between } \frac{\rho_{\mu s}\rho_{ms}}{1 \pm d}, \\ \rho^2_{\mu m|s} > \rho^2_{\mu m} & \text{if} & \rho_{\mu m} & \text{is between } \rho_{\mu s}\rho_{ms} \pm d \text{ but not between } \frac{\rho_{\mu s}\rho_{ms}}{1 \pm d}, \end{array}$$

where $d = \left[(1-\rho^2_{\mu s})(1-\rho^2_{ms})\right]^{\frac{1}{2}}$.

Proposition 4 suggests that if we increase the number of portfolios, the explanatory power of an asset pricing model within the portfolios could be eventually higher or lower than that for the firms in the entire population. Which one is the case crucially depends on $\rho_{\mu m}$, $\rho_{\mu s}$, and ρ_{ms}. Simply because μ and s are highly correlated does not guarantee that an asset pricing model will do a worse job within portfolios than for the entire population. In Figure 1, we plot the feasible regions of $\rho_{\mu m}$ and ρ_{ms} when $\rho_{\mu s} = 0.9$. The grey region is the region where $\rho^2_{\mu m} \geq \rho^2_{\mu m|s}$ and the dark region is the region where $\rho^2_{\mu m} < \rho^2_{\mu m|s}$. Even though μ and s are highly correlated with each other in this case, we can still see that there is quite a wide range of $\rho_{\mu m}$ and ρ_{ms} that will lead to $\rho^2_{\mu m} < \rho^2_{\mu m|s}$. Without knowing the exact values of $\rho_{\mu m}$, $\rho_{\mu s}$, and ρ_{ms}, it is difficult to conclude that sorting firms into a large number of portfolios will make an asset pricing model looks bad within the portfolios.

II. Simulation Evidence

In the previous section, we derive the theoretical explanatory power of an asset pricing model across and within portfolios. However, the measures that we derive are for the population. In empirical studies, we can only use finite number of firms in our sample, so it raises the question of how relevant our theoretical results are in practice. We address this question by simulation. The basic design of our simulation experiment is as follows. We sample 9000 firms from a population that has a multivariate normal distribution of (μ, m, s). The size of our sample roughly corresponds to the total number of firms available in the combined NYSE, AMEX and NASDAQ at the end of 1997.

For the marginal distributions, we assume both μ and m have a mean of 1%/month and a standard deviation of 0.2%/month. Under this assumption, 99% of the firms have expected return and predicted expected return in the range of 0.48%/month to 1.52%/month.[19] The sorting variable s has a mean of 6 and a standard deviation of 1.5. We interpret s as the natural logarithm of the market capitalization (in million dollars) of the equity of the firm. With this assumption, 99% of the firms will have market capitalization falls between the range of \$8.5 million to \$19.2 trillion. We have two sets of simulations. In both sets of simulations, we assume $\rho_{\mu s} = -0.8$ and $\rho_{ms} = -0.5$, so the expected return and the predicted expected return are assumed to be negatively correlated with the sorting variable. The choice of negative values for $\rho_{\mu s}$ and ρ_{ms} is motivated by the observations that small firms seem to have higher average returns as well as higher market betas. The only difference between the two sets of simulations is that we assume $\rho_{\mu m} = 0$ for the first set of simulations and $\rho_{\mu m} = 0.8$ for the second set. Therefore, in the first set of simulations, the asset pricing model alone is completely incapable of explaining the expected returns of individual firms whereas in the second set of simulations, the asset pricing model provides strong explanatory power on the expected returns on individual firms. In Figures 2 and 3, we plot the variables against each other for each set of simulations. Since the 9000 firms are only a sample from the population, the sample correlation coefficients are close but not exactly the same as the population ones.

For each set of simulations, we sort the firms into $n = 10$, 25, 50, and 100 portfolios based on s, each with an equal number of firms. We then form equally weighted as well as value-weighted portfolios (assuming s is the natural logarithm of size) to examine the explanatory power of the asset pricing model across these portfolios. In Panel A of Table I, we present the sample explanatory power of the asset pricing model, $R^2_{\mu m}(n)$, across different set of portfolios. To help us to see the relation between μ^i_p and m^i_p of these portfolios, we also provide a scatter plot of μ^i_p against m^i_p for the cases of 10 and

[19] Alternatively, by dropping the percentage per month, we can treat m as the CAPM beta.

100 equally weighted and value-weighted portfolios in Figures 4 and 5.

Table I about here

From both the plots as well as Panel A of Table I, we can observe that regardless of equally weighted or value-weighted portfolios, the asset pricing model provides very strong explanatory power across all sets of portfolios. Although the asset pricing model is completely incapable of explaining expected returns on individual firms for the first set of simulations, its deficiency is hardly detectable at the portfolio level. Compared with the very good asset pricing model in the second set of simulation, the $R^2_{\mu m}(n)$ at the portfolio level only display minor differences. For a researcher looking at these portfolios, it is difficult for him to distinguish a good model from a bad model. As n increases, $R^2_{\mu m}(n)$ moves further away from its theoretical value of one. This is because when n is large, the number of firms in each portfolio will not be large enough for the average pricing errors to converge to zero. However, even for $n = 100$, the asset pricing model still explains more than 93% of the cross-sectional variations of the expected returns at the portfolio level. With this kind of consistent performance of the asset pricing model across different sets of portfolios, one would be easily tempted to conclude that the asset pricing model is a good model. Unfortunately, the truth is $R^2_{\mu m}(n)$ at the portfolio level tells us absolutely nothing about $\rho^2_{\mu m}$ at the individual firm level.

In Panel B of Table I, we present the sample explanatory power of the two asset pricing models within the n portfolios, $R^2_{\mu m \cdot i}$. Since there are many portfolios, instead of presenting $R^2_{\mu m \cdot i}$ for all the portfolios, we simply present the minimum, maximum and average $R^2_{\mu m \cdot i}$ of the n portfolios. For the first set of simulations, we can see that the asset pricing model performs substantially better within the portfolios than for the whole sample. For all the cases, we find that the average $R^2_{\mu m \cdot i}$ is always greater than 50%. From the within-portfolio explanatory power numbers, it would be difficult for one to infer that the asset pricing model is in fact completely incapable of explaining expected returns at the individual firm level. For the second set of simulations, the numbers for the within-portfolio explanatory power of the asset pricing model are mostly lower than that for the population ($\rho^2_{\mu m} = 0.64$). However, the average $R^2_{\mu m \cdot i}$ stays pretty much the same as n increases, showing no sign of converging to zero.[20] Although $\rho_{\mu m}$ plays a role in determining $\rho^2_{\mu m|s}$, we still could not use $R^2_{\mu m \cdot i}$ to obtain meaningful inference about $\rho^2_{\mu m}$. For example, even though the $\rho^2_{\mu m}$ are very different for the two sets of simulation, the average $R^2_{\mu m \cdot i}$ for the first set of simulations is about the same or larger than those for the second set of simulations. Therefore, just like the sample explanatory power of an asset pricing model across portfolios, the sample explanatory power of an asset pricing model within portfolios cannot be used to evaluate asset pricing models.

III. Concluding Remarks

The message in this paper is clear. It is that the explanatory power of an asset pricing model across and within portfolios both do not provide useful information about the explanatory power of the asset pricing model at the individual firm level. For the across portfolios case, the only thing that matters is the sorting variable. The explanatory power of the asset pricing model is only a function of the sorting variable and its relation with the expected return and the predicted expected return by the asset pricing model. The joint distribution of the expected return and the predicted expected return plays no role in determining the explanatory power of the asset pricing model at the portfolio level. For the within portfolio case, while the correlation between expected return and predicted expected return plays a role in determining the explanatory power of the asset pricing model within portfolios, the correlations between the sorting variable and the expected and predicted expected returns also play important roles. Depending on the correlations between the three variables, the explanatory power of an asset pricing model

[20]The theoretical limit for the within-portfolio explanatory power in the second set of simulations is $\rho^2_{\mu m|s} = 0.593$.

at the individual firm level can be higher or lower than its explanatory power within the portfolios. Therefore, neither the range nor the average explanatory power of an asset pricing model within the portfolios are good indicators of the explanatory power of the asset pricing model at the individual firm level.

The implications of our results are disturbing. Taken at face value, one could infer that we learn nothing from empirical studies that use portfolios. One could also suggest that we should never use portfolios in empirical asset pricing studies since they provide no meaningful information. Our view is far less pessimistic. There are merits to using portfolios, and researchers will continue to use portfolios in empirical studies in the foreseeable future. While there are problems with using portfolios, there are equally many problems with using individual firms. The important task for finance researchers is to find ways to mitigate and address these problems.

While we do not have perfect solutions to these problems, we can make one suggestion. It is that we should not simply look at $\rho^2_{\mu m}(n)$ and $\rho^2_{\mu m \cdot i}$ as the only measures of explanatory power of an asset pricing model. For some theoretical asset pricing models, they make predictions about the magnitude of the coefficients in relating expected return to the risk measures. Imposing additional restrictions like this will help us to detect misspecifications even at the portfolio level. However, this will put empirically motivated models at an unfair advantage since they only suggest variables that are correlated with expected returns, without making predictions about the magnitude or even the sign of their slope coefficients.

Given the importance of the use of portfolios in empirical asset pricing studies, we hope future research will continue to address the important issues of how portfolios should be formed, and how we should properly interpret empirical results that use portfolios.

Appendix

Proof of Proposition 1: Substituting $E[\mu|s] = a + bE[m|s]$ in the definition of m_p^i, we obtain $\mu_p^i = a + bm_p^i$. As long as $b \neq 0$, μ_p^i is a linear function of m_p^i and we have $\rho_{\mu m}^2(n) = 1$. Otherwise, if $b = 0$, μ_p^i are constant across portfolios and we have $\rho_{\mu m}^2(n) = 0$. *Q.E.D.*

Proof of Proposition 2: We start off the proof by deriving the expression of $\sigma_{\mu m \cdot i}$ for the case of multivariate elliptical distribution. From (18), we have

$$\begin{aligned}
\sigma_{\mu m \cdot i} &= n \int_{s_{i-1}^*}^{s_i^*} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mu - \mu_p^i)(m - m_p^i) f_y(\mu, m, s) \mathrm{d}\mu \mathrm{d}m \mathrm{d}s \\
&= \int_{s_{i-1}^*}^{s_i^*} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mu - E[\mu|s] + E[\mu|s] - \mu_p^i)(m - E[m|s] + E[m|s] - m_p^i) \right. \\
&\qquad \left. f_{\mu m|s}(\mu, m|s) \mathrm{d}\mu \mathrm{d}m \right] n f_s(s) \mathrm{d}s \\
&= \int_{s_{i-1}^*}^{s_i^*} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mu - E[\mu|s])(m - E[m|s]) f_{\mu m|s}(\mu, m|s) \mathrm{d}\mu \mathrm{d}m \right] n f_s(s) \mathrm{d}s \\
&\quad + \int_{s_{i-1}^*}^{s_i^*} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (\mu - E[\mu|s])(E[m|s] - m_p^i) f_{\mu m|s}(\mu, m|s) \mathrm{d}\mu \mathrm{d}m \right] n f_s(s) \mathrm{d}s \\
&\quad + \int_{s_{i-1}^*}^{s_i^*} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (E[\mu|s] - \mu_p^i)(m - E[m|s]) f_{\mu m|s}(\mu, m|s) \mathrm{d}\mu \mathrm{d}m \right] n f_s(s) \mathrm{d}s \\
&\quad + \int_{s_{i-1}^*}^{s_i^*} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (E[\mu|s] - \mu_p^i)(E[m|s] - m_p^i) f_{\mu m|s}(\mu, m|s) \mathrm{d}\mu \mathrm{d}m \right] n f_s(s) \mathrm{d}s.
\end{aligned}$$

The expression inside the brackets of the first term is the conditional covariance between μ and m. Under the assumption that y has a multivariate elliptical distribution of, it is equal to (see, for example, Theorem 1.5.4 of Muirhead (1982))

$$g(s) \left[\sigma_{\mu m} - \frac{\sigma_{\mu s} \sigma_{m s}}{\sigma_s^2} \right],$$

for some positive function $g(s)$. The second term is equal to zero because $E[m|s] - m_p^i$ is not a function of μ and m, so it can be taken out of the brackets. Similarly, the third term is also equal to zero. For the last term, we take $(E[\mu|s] - \mu_p^i)(E[m|s] - m_p^i)$ outside of the brackets and it becomes

$$\int_{s_{i-1}^*}^{s_i^*} (E[\mu|s] - \mu_p^i)(E[m|s] - m_p^i) n f_s(s) \mathrm{d}s.$$

Under the assumption of multivariate elliptical distribution, we have

$$\begin{aligned} E[\mu|s] &= E[\mu] + \frac{\sigma_{\mu s}}{\sigma_s^2}(s - E[s]) \\ &= \left(E[\mu] - \frac{\sigma_{\mu s}}{\sigma_s^2} E[s]\right) + \frac{\sigma_{\mu s}}{\sigma_s^2} s. \end{aligned} \tag{A1}$$

Substituting (A1) in (5), we have

$$\begin{aligned} \mu_p^i &= n \int_{s_{i-1}^*}^{s_i^*} \left[\left(E[\mu] - \frac{\sigma_{\mu s}}{\sigma_s^2} E[s]\right) + \frac{\sigma_{\mu s}}{\sigma_s^2} s\right] f_s(s)\mathrm{d}s \\ &= \left(E[\mu] - \frac{\sigma_{\mu s}}{\sigma_s^2} E[s]\right) + \frac{\sigma_{\mu s}}{\sigma_s^2} s_p^i, \end{aligned} \tag{A2}$$

where the second equality follows from the fact that

$$\int_{s_{i-1}^*}^{s_i^*} f_s(s)\mathrm{d}s = \frac{1}{n}. \tag{A3}$$

From (A1) and (A2), we have

$$E[\mu|s] - \mu_p^i = \frac{\sigma_{\mu s}}{\sigma_s^2}(s - s_p^i).$$

Similarly,

$$E[m|s] - m_p^i = \frac{\sigma_{ms}}{\sigma_s^2}(s - s_p^i).$$

Substituting these expressions back, we have

$$\begin{aligned} \sigma_{\mu m \cdot i} &= \left(\sigma_{\mu m} - \frac{\sigma_{\mu s}\sigma_{ms}}{\sigma_s^2}\right) \int_{s_{i-1}^*}^{s_i^*} n g(s) f_s(s)\mathrm{d}s + \left(\frac{\sigma_{\mu s}\sigma_{ms}}{\sigma_s^2}\right) \int_{s_{i-1}^*}^{s_i^*} n \left(\frac{s - s_p^i}{\sigma_s}\right)^2 f_s(s)\mathrm{d}s \\ &= \left(\int_{s_{i-1}^*}^{s_i^*} n g(s) f_s(s)\mathrm{d}s\right) \left[\sigma_{\mu m} - \left(\frac{\sigma_{\mu s}\sigma_{ms}}{\sigma_s^2}\right) h(i)\right]. \end{aligned} \tag{A4}$$

Using a similar proof, we can show that

$$\sigma_{\mu \cdot i}^2 = \left(\int_{s_{i-1}^*}^{s_i^*} n g(s) f_s(s)\mathrm{d}s\right) \left[\sigma_\mu^2 - \left(\frac{\sigma_{\mu s}^2}{\sigma_s^2}\right) h(i)\right], \tag{A5}$$

$$\sigma_{m \cdot i}^2 = \left(\int_{s_{i-1}^*}^{s_i^*} n g(s) f_s(s)\mathrm{d}s\right) \left[\sigma_m^2 - \left(\frac{\sigma_{ms}^2}{\sigma_s^2}\right) h(i)\right]. \tag{A6}$$

With these expressions, $\rho_{\mu m \cdot i}^2$ follows trivially from the definition of correlation coefficient. For the case of multivariate normality, $g(s) = 1$ and hence the denominator of

$h(i) = 1$ from (A3). The integral in the numerator is the variance of a doubly truncated normal distribution. Using (13.135) of Johnson and Kotz (1994), we have

$$\int_{s_{i-1}^*}^{s_i^*} n\left(\frac{s-s_p^i}{\sigma_s}\right)^2 f_s(s)\mathrm{d}s = 1 - n^2\left[\phi(c_{i-1}^*) - \phi(c_i^*)\right]^2 + n\left[c_{i-1}^*\phi(c_{i-1}^*) - c_i^*\phi(c_i^*)\right]. \quad \text{(A7)}$$

This completes the proof. *Q.E.D.*

Proof of Proposition 3: Without loss of generality, we assume $E[s] = 0$ and $\sigma_s = 1$. For any $\epsilon_2 > 0$, we can find an $M > 0$ such that $P[-M < s < M] > 1 - \epsilon_2$. Therefore, allowing for possible rounding, there will be at least $n(1-\epsilon_2)-2$ portfolios that consists of firms with $s \in (-M, M)$. Define $f^* = \min_{-M\le s\le M} f_s(s)$ and $g^* = \min_{-M\le s\le M} g(s)$. Let $n_0 = \frac{1}{f^*\sqrt{g^*\epsilon_1}}$, then for $n > n_0$, we have

$$(s_i^* - s_{i-1}^*)f^* \le \frac{1}{n} < \frac{1}{n_0} = f^*\sqrt{g^*\epsilon_1},$$

and hence $(s_i^* - s_{i-1}^*)^2 < g^*\epsilon_1$. When $n > n_0$, then for every portfolio that consists of firms with $s \in (-M, M)$, we have

$$\begin{aligned} h(i) &= 1 - \frac{\int_{s_{i-1}^*}^{s_i^*} n(s-s_p^i)^2 f_s(s)\mathrm{d}s}{\int_{s_{i-1}^*}^{s_i^*} ng(s)f_s(s)\mathrm{d}s} \\ &> 1 - \frac{\int_{s_{i-1}^*}^{s_i^*} n(s_i^*-s_{i-1}^*)^2 f_s(s)\mathrm{d}s}{\int_{s_{i-1}^*}^{s_i^*} ng^* f_s(s)\mathrm{d}s} \\ &> 1 - \frac{\int_{s_{i-1}^*}^{s_i^*} ng^*\epsilon_1 f_s(s)\mathrm{d}s}{\int_{s_{i-1}^*}^{s_i^*} ng^* f_s(s)\mathrm{d}s} \\ &= 1 - \epsilon_1. \end{aligned}$$

For the case of normality, we need to show the portfolios of firms with $s \notin (-M, M)$ also have $h(i) \to 1$. By symmetry, we just need to show the portfolios with $s \in (-\infty, -M]$ will have $h(i) \to 1$. Denote $v_1, v_2, \ldots, v_k$ the variance of s for the first to the kth portfolio that consists of firms with $s \in (-\infty, -M]$. For the case of normality, $g(s) = 1$ and $h(i) = 1 - v_i$, so we just need to show $\lim_{n\to\infty} v_i = 0$ for this set of portfolios. Since the normal density is increasing over $(-\infty, -M]$, it is easy to show that $v_1 \ge v_2 \cdots \ge v_k$, and hence it suffices to show $\lim_{n\to\infty} v_1 = 0$. For the first portfolio, denote $\Phi(\cdot)$ as the cumulative density function of the standard normal distribution, we have

$$\begin{aligned}
\lim_{n\to\infty} v_1 &= \lim_{c_1^*\to-\infty} 1 - \frac{c_1^*\phi(c_1^*)}{\Phi(c_1^*)} - \left[\frac{\phi(c_1^*)}{\Phi(c_1^*)}\right]^2 \\
&= \lim_{c_1^*\to-\infty} \frac{\Phi(c_1^*)^2 - c_1^*\phi(c_1^*)\Phi(c_1^*) - \phi(c_1^*)^2}{\Phi(c_1^*)^2} \\
&= \lim_{c_1^*\to-\infty} \frac{1}{(c_1^*)^2 - 1} = 0,
\end{aligned}$$

where the last line is obtained by repeated use of L'Hôpital's rule. This completes the proof. *Q.E.D.*

Proof of Proposition 4: We first establish the feasible range of $\rho_{\mu m}$ for given values of $\rho_{\mu s}$ and ρ_{ms}. It is well known that the necessary and sufficient condition for a matrix to be positive semidefinite is that all of its principal minors are nonnegative. For a 3 by 3 correlation matrix, this condition is equivalent to requiring its determinant to be nonnegative. The determinant of the correlation matrix of y is given by

$$-\rho_{\mu m}^2 + 2\rho_{\mu s}\rho_{ms}\rho_{\mu m} + (1 - \rho_{\mu s}^2 - \rho_{ms}^2), \tag{A8}$$

which is a quadratic equation in $\rho_{\mu m}$. For it to be nonnegative, $\rho_{\mu m}$ must lie between the roots of the quadratic equation, which are

$$\rho_{\mu s}\rho_{ms} \pm \left(\rho_{\mu s}^2\rho_{ms}^2 + 1 - \rho_{\mu s}^2 - \rho_{ms}^2\right)^{\frac{1}{2}} = \rho_{\mu s}\rho_{ms} \pm d. \tag{A9}$$

For $\rho_{\mu m}^2 \geq \rho_{\mu m|s}^2$, we need

$$\begin{aligned}
& d^2\rho_{\mu m}^2 - (\rho_{\mu m} - \rho_{\mu s}\rho_{ms})^2 \geq 0 \\
\Rightarrow\ & (d^2 - 1)\rho_{\mu m}^2 + 2\rho_{\mu s}\rho_{ms}\rho_{\mu m} - \rho_{\mu s}^2\rho_{ms}^2 \geq 0.
\end{aligned} \tag{A10}$$

The left hand side of the equation is also a quadratic equation in $\rho_{\mu m}$ and it will be nonnegative if $\rho_{\mu m}$ is between the two roots:

$$\frac{-\rho_{\mu s}\rho_{ms} \pm \rho_{\mu s}\rho_{ms} d}{d^2 - 1} = \frac{\rho_{\mu s}\rho_{ms}}{1 \pm d}. \tag{A11}$$

Since $0 \leq d \leq 1$, these two roots fall inside the feasible range of $\rho_{\mu m}$. This completes the proof. *Q.E.D.*

References

Berk, Jonathan B., 1998, Sorting out sorts, *Journal of Finance* forthcoming.

Chen, Naifu, Raymond Kan, and Chu Zhang, 1998, A critique of the use of t-ratios in model selection, Working paper, University of Toronto.

Grauer, Robert R., and Johannus A. Janmaat, 1998, The unintended consequences of grouping in tests of asset pricing models, Working paper, Simon Fraser University.

Johnson, Norman L., Samuel Kotz, and N. Balakrishnan, 1994, *Continuous Univariate Distributions, Volume 1* (Wiley, New York).

Kan, Raymond, and Chu Zhang, 1995, The value-weighted average misspecification in asset pricing models, Working paper, University of Toronto.

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

Kelker, Douglas, 1970, Distribution theory of spherical distributions and a location-scale parameter generalization, *Sankhyā: The Indian Journal of Statistics* 32, Series A, 419–430.

Lo, Andrew W., and A. Craig MacKinlay, 1990, Data-snooping biases in tests of financial asset pricing models, *Review of Financial Studies* 3, 431–467.

Muirhead, Robb J., 1982, *Aspects of Multivariate Statistical Theory* (Wiley, New York).

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

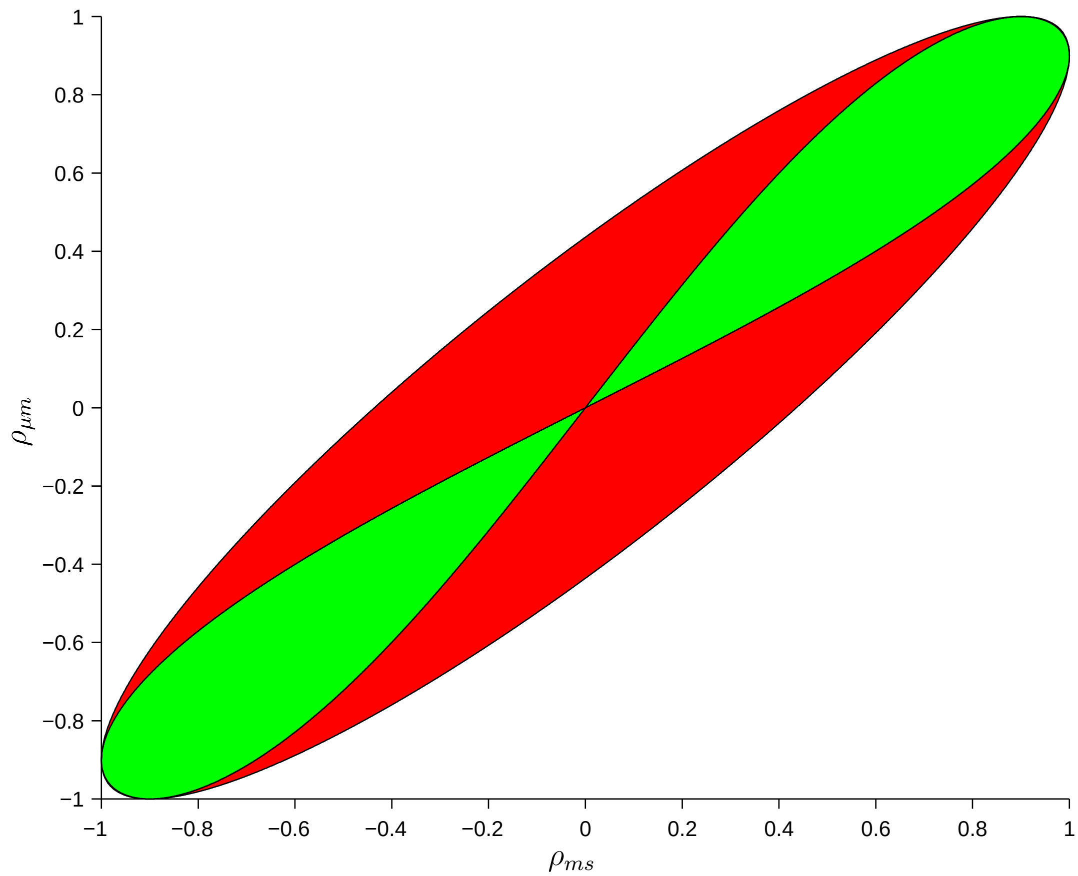

Figure 1: Feasible Regions of $\rho_{\mu m}$ and ρ_{ms} when $\rho_{\mu s} = 0.9$

The figure presents the feasible regions of ρ_{ms} and $\rho_{\mu m}$ when $\rho_{\mu s} = 0.9$. The grey region shows the combinations of $\rho_{\mu m}$ and $\rho_{\mu s}$ that leads to $\rho^2_{\mu m|s} \leq \rho^2_{\mu m}$, and the dark region shows the combinations of $\rho_{\mu m}$ and $\rho_{\mu s}$ that leads to $\rho^2_{\mu m|s} > \rho^2_{\mu m}$. μ is the expected return of a firm, m is its predicted expected return by an asset pricing model, and s is the value of a firm-specific variable which is used to sort firms into portfolios. ρ is used to denote correlation coefficients between different pairs of variables, and $\rho^2_{\mu m|s}$ is the partial coefficient of determination between μ and m, conditional on the firm-specific variable is equal to s. $\rho^2_{\mu m|s}$ is a measure of explanatory power of an asset pricing within a portfolio, when the number of portfolios increases to infinity.

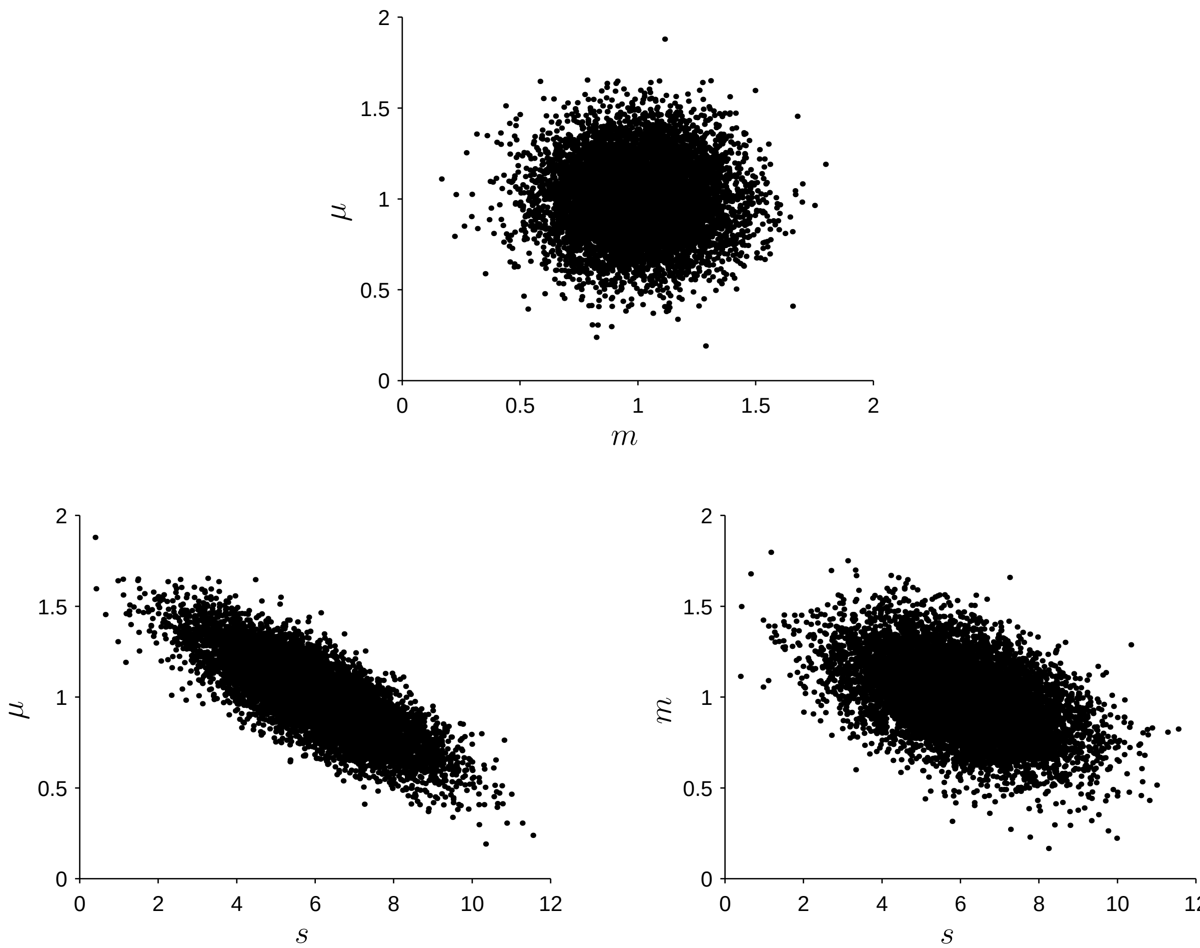

Figure 2: Scatter Plots of Expected Returns, Predicted Expected Returns and Sorting Variables of 9000 Firms for the First Set of Simulations

The figure presents scatter plots of expected return (μ) vs. predicted expected return (m), expected return vs. sorting variable (s), and predicted expected return vs. sorting variable (s) for a sample of 9000 firms. The firms are drawn from a population where (μ, m, s) have a multivariate normal distribution with $\rho_{\mu m} = 0$, $\rho_{\mu s} = -0.8$, and $\rho_{ms} = -0.5$.

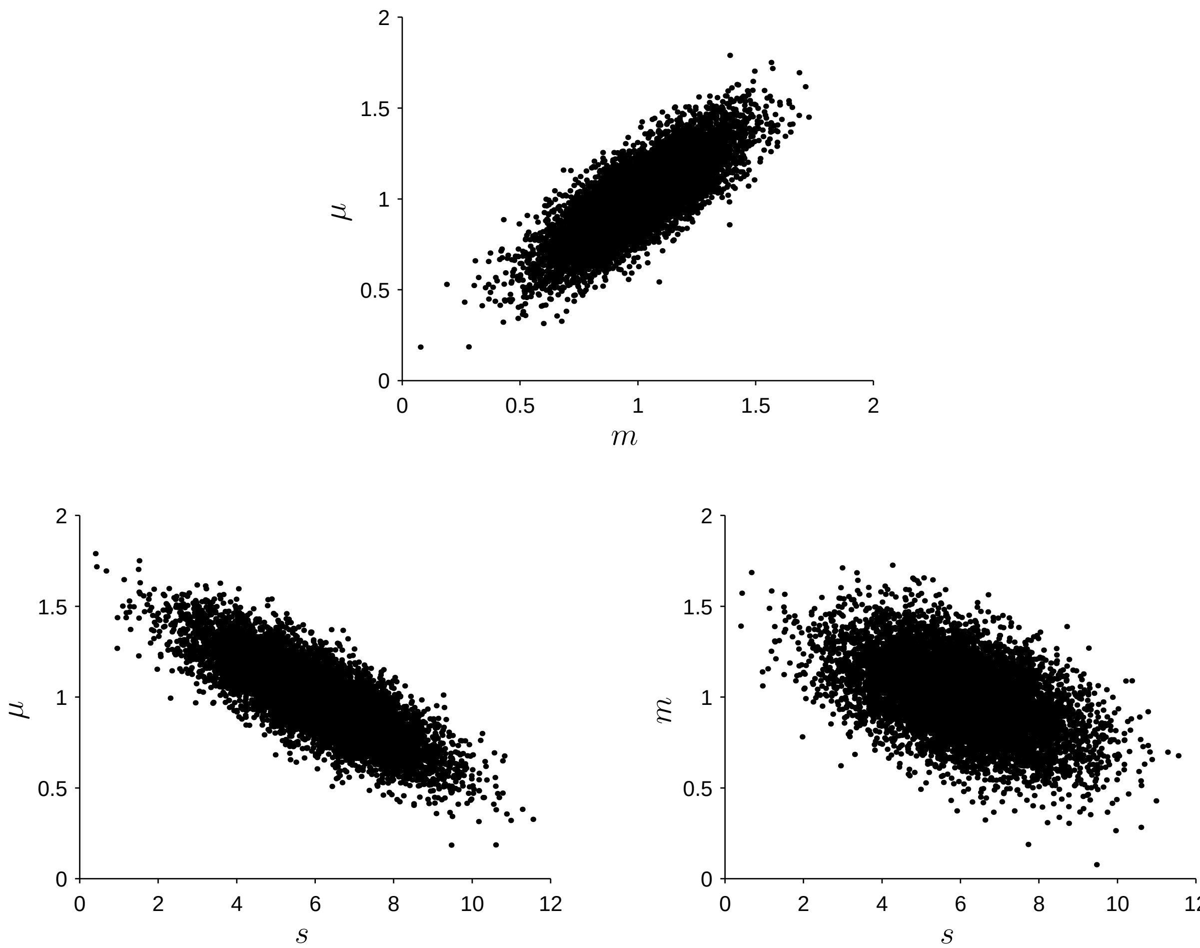

Figure 3: Scatter Plots of Expected Returns, Predicted Expected Returns and Sorting Variables of 9000 Firms for the Second Set of Simulations

The figure presents scatter plots of expected return (μ) vs. predicted expected return (m), expected return vs. sorting variable (s), and predicted expected return vs. sorting variable for a sample of 9000 firms. The firms are drawn from a population where (μ, m, s) have a multivariate normal distribution with $\rho_{\mu m} = 0.8$, $\rho_{\mu s} = -0.8$, and $\rho_{ms} = -0.5$.

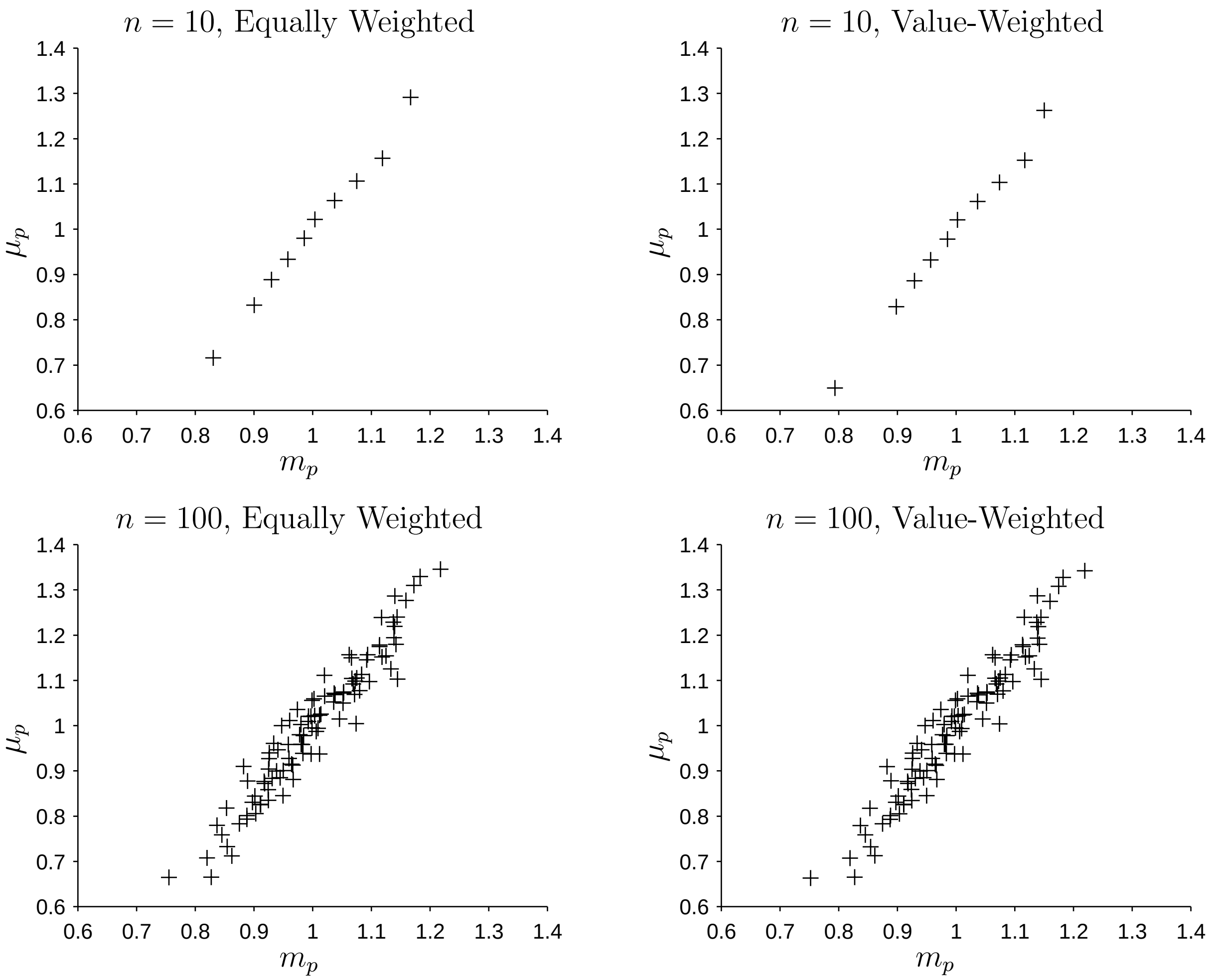

Figure 4: Scatter Plots of Expected Returns and Predicted Expected Returns for 10 and 100 Equally Weighted and Value-Weighted Portfolios for the First Set of Simulations

The figure presents scatter plots of expected portfolio return (μ_p) vs. predicted expected portfolio return (m_p), of 10 and 100 equally weighted and value-weighted portfolios. The portfolios are formed based on 9000 firms which are drawn from a population where the expected return (μ), the predicted expected return (m), and the sorting variable of the individual firms have a multivariate normal distribution with $\rho_{\mu m} = 0$, $\rho_{\mu s} = -0.8$, and $\rho_{ms} = -0.5$. The 9000 firms are sorted into 10 and 100 portfolios based on s, and each portfolio has the same number of firms.

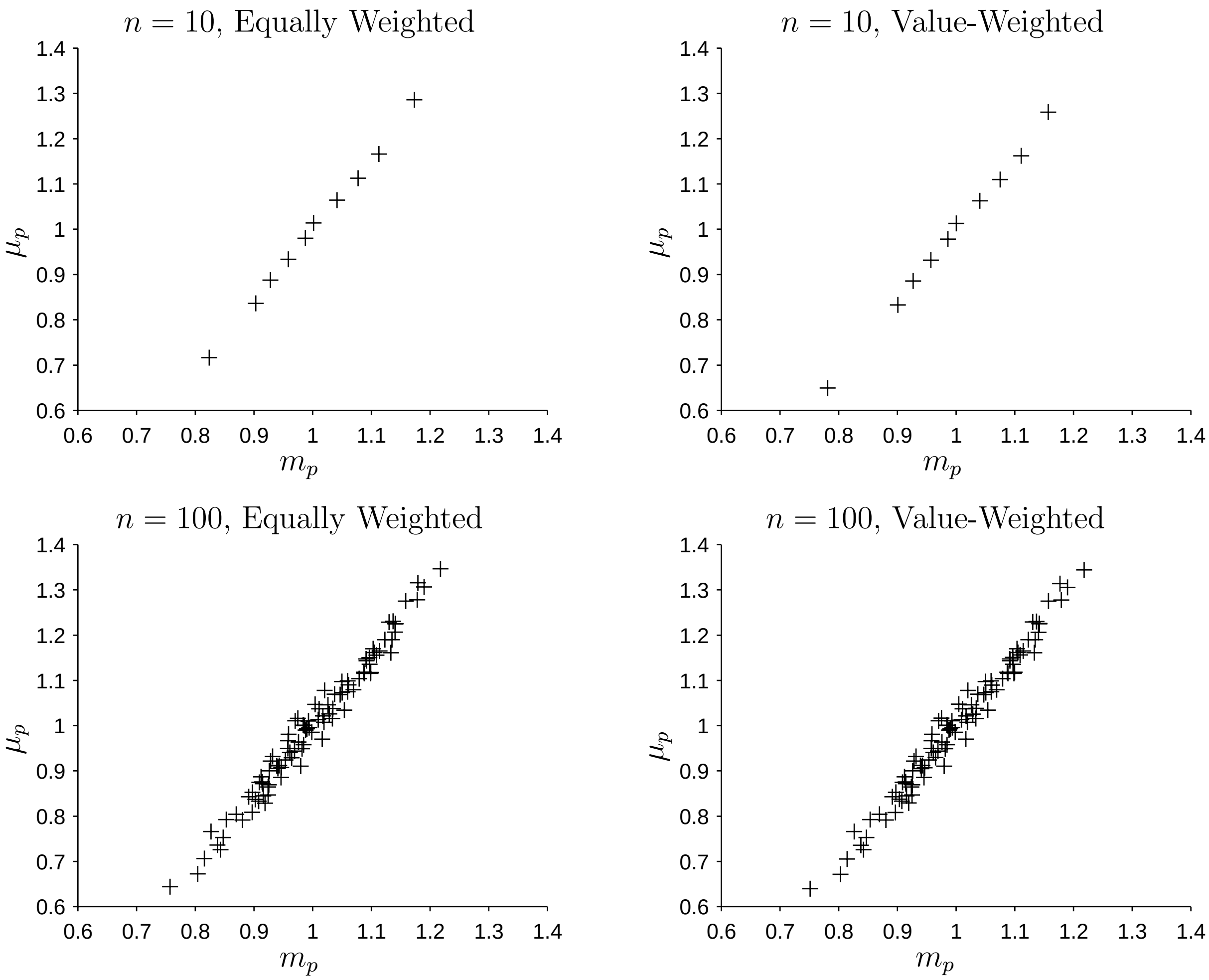

Figure 5: Scatter Plots of Expected Returns and Predicted Expected Returns for 10 and 100 Equally Weighted and Value-Weighted Portfolios for the Second Set of Simulations

The figure presents scatter plots of expected portfolio return (μ_p) vs. predicted expected portfolio return (m_p), of 10 and 100 equally weighted and value-weighted portfolios. The portfolios are formed based on 9000 firms which are drawn from a population where the expected return (μ), the predicted expected return (m), and the sorting variable (s) of the individual firms have a multivariate normal distribution with $\rho_{\mu m} = 0.8$, $\rho_{\mu s} = -0.8$, and $\rho_{ms} = -0.5$. The 9000 firms are sorted into 10 and 100 portfolios based on s, and each portfolio has the same number of firms.

Table I
Explanatory Power of Asset Pricing Models Across and Within Portfolios

Panel A of the table presents the sample explanatory power of the asset pricing models across different sets of portfolios ($R^2_{\mu m}(n)$). The portfolios are formed based on 9000 firms which are drawn from a population where the expected return (μ), the predicted expected return (m), and the sorting variable (s) of individual firms have a multivariate normal distribution with $\rho_{\mu s} = -0.8$ and $\rho_{ms} = -0.5$. Two sets of results are reported, one under the assumption that $\rho_{\mu m} = 0$ and the other under the assumption that $\rho_{\mu m} = 0.8$. The 9000 firms are sorted into n portfolios based on s, and each portfolio has the same number of firms. We present the cases of $n = 10$, 20, 50, and 100. For each case, we present the sample explanatory power of the asset pricing model across the equally weighted as well as the value-weighted portfolios. Panel B of the table presents the sample explanatory power of the asset pricing models within different sets of portfolios ($R^2_{\mu m \cdot i}$) in Panel A. For each case, we present the minimum, the maximum, and the average sample explanatory power of the asset pricing models within the n portfolios.

Panel A: Explanatory Power of Asset Pricing Models Across Portfolios

	$R^2_{\mu m}(n)$			
	$\rho_{\mu m} = 0$		$\rho_{\mu m} = 0.8$	
n	Equally Weighted	Value-Weighted	Equally Weighted	Value-Weighted
10	0.991	0.991	0.998	0.998
20	0.988	0.988	0.997	0.997
50	0.965	0.966	0.991	0.991
100	0.931	0.930	0.978	0.978

Panel B: Explanatory Power of Asset Pricing Models Within Portfolios

	Distribution of $R^2_{\mu m \cdot i}$					
	$\rho_{\mu m} = 0$			$\rho_{\mu m} = 0.8$		
n	Minimum	Maximum	Average	Minimum	Maximum	Average
10	0.272	0.620	0.521	0.541	0.622	0.584
20	0.318	0.647	0.565	0.510	0.627	0.584
50	0.279	0.728	0.590	0.446	0.675	0.583
100	0.289	0.756	0.595	0.357	0.736	0.584