DIVISION OF THE HUMANITIES AND SOCIAL SCIENCES

CALIFORNIA INSTITUTE OF TECHNOLOGY

PASADENA, CALIFORNIA 91125

TESTING THE MEAN-VARIANCE EFFICIENCY OF WELL-DIVERSIFIED PORTFOLIOS IN VERY LARGE CROSS-SECTIONS

Peter Bossaerts
California Institute of Technology

Pierre Hillion
INSEAD, France

SOCIAL SCIENCE WORKING PAPER 854

August 1993

Testing the Mean-Variance Efficiency of Well-Diversified Portfolios in Very Large Cross-Sections

Peter Bossaerts Pierre Hillion

Abstract

We propose a new way of testing the mean–variance efficiency of well-diversified portfolios that exploits the cross–sectional size of typical financial datasets. The methodology consists of a sequence of simple tests, the results of which are aggregated in a statistic. This statistic is shown to be asymptotically standard normally distributed, despite dependence, in cross–section and over time, of the idiosyncratic risk. We investigate theoretically the asymptotic power of our test against the alternative that the well–diversified portfolio is not mean-variance efficient. By construction, our procedure is more powerful than standard tests of mean–variance efficiency that combine the assets in the cross-section into a limited set of (arguably) arbitrary portfolios. Even in cases where the latter has zero power, it can have unit asymptotic power. The incremental power is evidenced in tests of the mean–variance efficiency of the value weighted portfolio of common stock listed on the NYSE and AMEX. Unlike previously thought, however, the selection bias caused by including only continuously traded securities in the test is found to be important. By running the test in a case where it is known to have zero power, we are able to empirically confirm the correctness of the theoretical asymptotic properties of our statistic.

Keywords: Mean–Variance Efficiency, Well–diversified Portfolios, Exchangeable Random Variables, De Finetti's Theorem, Central Limit Theorem, Law of the Iterated Logarithm.

Testing the Mean-Variance Efficiency of Well-Diversified Portfolios in Very Large Cross-Sections

Peter Bossaerts * Pierre Hillion †

1 Introduction

Traditionally, tests of the mean–variance efficiency of a portfolio combine assets into a limited number (five to ten) of portfolios. The returns on these portfolios in excess of the riskfree rate are regressed onto that of the candidate portfolio, and one tests whether the intercepts equal zero. While the early tests justified the portfolio arrangement technique by pointing to errors in variables in the estimation of the relevant parameters (see *e.g.* Black, Jensen and Scholes [1972], Fama and MacBeth [1973])[1], later work needed the reduction in the size of the cross–section in order to test *joint* restrictions on the intercepts (Gibbons [1982]). The multivariate tests in later studies are based on nonsingular estimates of the variance–covariance matrix of the regression errors. Nonsingularity requires the size of the cross–section be less than the length of the time series. The latter, however, is limited to at most five to ten years, because of stationarity considerations (Gibbons and Shanken [1987]).

Traditional tests of mean–variance efficiency have well–known disadvantages. Foremost, there is the lack of power (MacKinlay [1987], Gibbons, Ross and Shanken [1989]). As Heston [1992] explains, this is essentially due to the fact that absence of simple arbitrage opportunities restricts the intercepts in the regression of the excess returns of any reasonably diversified portfolio onto that of the candidate portfolio to be zero even if the

*HSS 228-77, California Institute of Technology, Pasadena, CA 91125 (U.S.A.), Phone +1.818.395.4028; e-mail: pbs@rioja.caltech.edu

†INSEAD, F-77305 Fontainebleau, France; Phone (1) 6072 4214; e-mail: hillion@freiba51.BITNET.
The authors thank participants at a seminar at Carnegie Mellon University and U.C. Riverside for comments. We thank Jet Propulsion Laboratory for access to their CRAY.

[1]For a comprehensive econometric analysis, see Shanken [1992].

latter is not mean–variance efficient. More recently, the portfolio formation technique has been criticized for a different reason: if mean–variance efficiency cannot be rejected on one portfolio–formation criterion (*e.g.*, based on the assets' "beta"), it may be when the assets are reshuffled into different portfolios (*e.g.*, based on company size); yet, it is not clear how to interpret the latter rejection, since it is based on the same data (Lo and MacKinlay [1990]).

This paper introduces a new way to test for mean–variance efficiency. It does not depend on portfolio building, and, hence, avoids some of the disadvantages of the standard testing procedure. The test uses a simple statistic that is computed from intercepts estimated in regressions of the excess return on individual assets in a cross–section onto that of the candidate mean–variance efficient portfolio. Keeping the length of the time series fixed, the statistic becomes normally distributed as the size of the cross–section increases. No distributional assumptions (such as normality) are imposed: we merely exploit the size of the cross–section. By construction, our test is more powerful than the traditional one. In particular, it can have unit (asymptotic) power even if the latter has zero power.[2] Our procedure also allows us to better gauge the effect of survivorship bias. Recent theoretical work (Brown, Goetzmann, Ibbotson and Ross [1992, p. 576]) has indicated that it may previously have been underestimated.

We analyze the properties of our test in the following context. We assume that there exists a riskfree asset.[3] Second, we assume that the candidate mean–variance efficient portfolio becomes well–diversified (in the sense of Chamberlain [1983]) as the size of the cross–section increases. Third, we assume that the regression errors become idiosyncratic, again as the cross–sectional size increases. In the terminology of Chamberlain [1983] and Chamberlain and Rothschild [1983], we thus assume that there is an approximate one–factor structure. When we analyze the power of our test against the alternative hypothesis that the candidate portfolio is not mean–variance efficient, we retain the assumption of diversification and the idiosyncratic nature of the regression errors.

We derive the asymptotic behavior of our test under the following statistical assumptions regarding the regression errors (idiosyncratic risk): we allow them to be dependent, in cross–section and over time, but impose the restriction that the joint distribution of any finite collection of them does not depend on the order. Technically, we assume that the regression errors are exchangeable. While this imposes cross–sectional homogeneity, it avoids independence assumptions. Exchangeability is a natural assumption in an environment where data limitations make it impossible to identify the nature of individual draws, yet one knows that sampling is not random. In our context, station-

[2]A test has unit asymptotic power if the probability of rejecting a false null hypothesis increases to one as the sample size increases. It has zero power if the probability of rejecting a false null hypothesis equals the probability of rejecting a correct null hypothesis.

[3]Hence, in the context of the Capital Asset Pricing Model, we would effectively test the Sharpe–Lintner model.

arity imposes limits on the time series length that preclude our detecting heterogeneity comprehensively.

In other words, the cross-sectional homogeneity implied by our assumption of exchangeable idiosyncratic risk has to be evaluated *against the background of nonstationarities in financial return series.* The lack of knowledge of the precise form of nonstationarity makes it impossible (or makes us unwilling) to use the times series dimension of our dataset to detect heterogeneity. Consequently, we are forced to consider all idiosyncratic risk as coming from a large, homogeneous pool (population).[4] Instead, we fully exploit the cross-sectional dimension of the dataset. Our procedure can thus be interpreted as providing a way to *learn about the nature of the mean-variance efficient frontier despite nonstationarities of unknown form* in the data.

We could have appealed to concepts such as strong mixing to derive the asymptotic properties of our test (as in, *e.g.*, Connor and Korajczyk [1992]). We choose not to do so, for two reasons. First, strong mixing is essentially a time series concept, depending on a natural order, which is absent in most cross-sections. Second, it would require regression errors sufficiently far apart in the cross-section not only to be uncorrelated, but also to be *independent.* Yet, there is evidence that squared errors satisfy a nontrivial factor structure (see, *e.g.*, Schwert and Seguin [1990, Table IV]), indicating that independence is unrealistic.[5] Exchangeability, in contrast, is a cross-sectional concept, and does not impose independence. Specifically, the squared regression errors need not be uncorrelated.

We apply our testing procedure to five-year periods of monthly returns on the value weighted portfolio of NYSE and AMEX stocks between 1926 and 1990, and confirm the incremental power of our procedure relative to the standard test using size-based portfolios. The rejections of mean-variance efficieny that we uncover may have been caused by arbitrage opportunities in the sense of Ross [1976]. We use the law of the iterated logarithm to shed some light on this possibility. When we reject mean-variance efficiency, we seem to find that this may be due to the presence of such arbitrage opportunities. In other words, we seem to reject Ross' Arbitrage Pricing Theory as well. We qualify this conclusion, however, by arguing that the cross-sectional size may still be too small to rely on the law of the iterated logarithm.

We also test the mean-variance efficiency of the *equally weighted* portfolio of NYSE and AMEX stock on the complete cross-section. In that case, we know that our test has zero power, and, hence, that our statistic should be distributed as if the null hypothesis were correct (provided the remaining assumptions are also satisfied). The values of our test statistic that we obtain over the thirteen subperiods confirm the asymptotic

[4]Exchangeability has been used before to define the boundaries of populations. For an in-depth analysis, see, *e.g.* Lindley and Novick [1981].

[5]The factor structure on squared errors implies that turbulence affects all securities simultaneously, but not necessarily in the same direction.

properties: we cannot reject that they are draws from a standard normal distribution. These results at the same time illustrate that the typical cross-sectional size of financial datasets is sufficient to rely on asymptotic statistical analysis.

Finally, our procedure provides an opportunity to investigate the effect of selection bias introduced by focusing exclusively on continuously traded securities. We find that the selection bias can be substantial, validating the theoretical analysis in Brown, Goetzmann, Ibbotson and Ross [1992].

The remainder of the paper is organized as follows. The next section introduces the test statistic and discusses its asymptotic properties. Section 3 elaborates on the power. Section 4 applies the methodology in tests of the mean-variance efficiency of the value weighted portfolio of NYSE and AMEX common stock. The correctness of the distributional properties of our statistic is proven empirically in Section 5. Section 6 investigates selection bias. Section 7 concludes by suggesting directions for further research.

2 Methodology

The test statistic is defined in this section, and its asymptotic properties are derived. Some of the material is very technical, and, hence, we will leave the reader the option to skip most of it and turn to Section 3 after a brief introduction.

We assume that there exists a riskfree asset, so that we can work with excess returns (returns minus the riskfree rate). Take a countably infinite sequence of assets, indexing the elements $i = 1, 2, 3, \ldots$. Let r_{it} denote the excess return on asset i over the tth period, where $t = 1, \ldots, T$. We keep T, the length of the time series, fixed. Consider the first N assets. N will be referred to as the size of the cross-section. Using the first N assets, we compute the value of a statistic that tests the mean-variance efficiency of our candidate portfolio. We are interested in knowing the properties of this statistic as $N \to \infty$.

Let $r_{p_N t}$ denote the excess return on the candidate portfolio. We will refer to it as 'portfolio p_N'. The composition of our candidate portfolio is allowed to change with N, hence the double-indexing. p_N should, however, have weights such that it becomes well-diversified as N increases. As in Chamberlain [1983], this is supposed to mean that the sum of the squared weights converges to zero.[6] Let p indicate the well-diversified limit portfolio.

Let ϵ_{iNt} be the error from projecting r_{it} onto $r_{p_N t}$, *i.e.*,

$$\epsilon_{iNt} = r_{it} - \alpha_{iN} - \beta_{iN} r_{p_N t} \quad (1)$$

[6]In the empirical application, we will consider two candidate portfolios: an equally weighted and a value weighted portfolio. While equal weighting guarantees eventual diversification, this can only be *assumed* in the case of value weighting.

with

$$E[\epsilon_{iNt}] = E[\epsilon_{iNt} r_{p_N t}] = 0.$$

As is customary in regression analysis, we shall condition on the regressor. In other words, we keep $r_{p_N t}$ "fixed in repeated sampling". In order to simplify the notation, however, we will not make this conditioning explicit. For instance, when we assume that the regression errors in (1) are exchangeable, we implicitly mean: conditional on $r_{p_N t}$.

The double–indexing of the intercept and the slope coefficient in Equation (1) reflects changes in the candidate portfolio as it becomes better diversified. More importantly, it is meant to capture the increasingly stringent restrictions that no–arbitrage conditions impose on the intercepts. This will be important in the analysis of the power of our testing procedure, to be discussed in Section 3.

We will be interested in the asymptotic properties of a statistic that tests whether p_N is mean–variance efficient (all N). Therefore, let us define the null hypothesis as:

$$\alpha_{iN} = 0, \tag{2}$$

all i and N, where α_{iN} is the intercept in (1).[7]

Some readers may not be interested in the technicalities of what follows. Therefore, let us briefly summarize the arguments, so that those who wish can immediately skip to the power analysis (Section 3) and the application (Section 4) without loss of continuity. Our test is based on estimates of the α_{iN}s in (1). It takes the cross–sectional average of these estimates, and scales it by the square root of the cross–sectional size ($\sqrt{N}$) and an estimate of the precision.[8] This statistic is denoted z_N. It is assumed that the regression errors are idiosyncratic (in the sense of Chamberlain [1983]) and exchangeable (cross–sectionally homogeneous but potentially dependent). Under these assumptions, we show that z_N satisfies a central limit theorem: its asymptotic distribution (as $N \to \infty$) is standard normal. The asymptotic distribution can be used in a test of the mean–variance efficiency of the candidate portfolio, in particular, whether $\alpha_{iN} = 0$, all i and N.

We will make use of central limit theorems that apply to the regression errors defined in (1). They are double–indexed as well, but this causes no problem, as standard central limit arguments are based on double–indexed sequences of random variables. Usually, however, the relationship between variables (indexed i) across sequences (indexed N) is left unrestricted, whereas we will need some minimal link. Part of this link is provided

[7]This null hypothesis is actually a bit stricter than what we would ideally aim at: we require *all* candidate portfolios p_N in the sequence to be mean–variance efficient in the respective cross–sections, instead of imposing only that they become mean–variance efficient as they converge to p.

[8]The reason for this rescaling should be clear: we rely on asymptotic distribution theory. If we do not rescale the average estimated intercept, its distribution would become degenerate as the cross–sectional size increases. Specifically, the rescaling with the square root of the sample size is precisely what one needs to obtain a nontrivial asymptotic distribution.

by the assumption that the regression errors become idiosyncratic as $N \rightarrow \infty$. Let us discuss this now.

We assume that the regression errors in (1) become idiosyncratic with the sample size. As in Chamberlain [1983], this means the following. Let Σ_{MN} be the matrix of covariances of the first M projection errors when the regressor is $r_{p_N t}$ (element (i,j) of this matrix equals $E[\epsilon_{iNt}\epsilon_{jNt}]$). If $M = N$, we write: Σ_{NN}. If A is any square matrix, let $\lambda(A)$ denote the set of its eigenvalues. We assume:

Assumption 1

$$\sup_N\{\sup\{\lambda(\Sigma_{NN})\}\} < \infty.$$

Assumption 1 immediately implies that the projection errors have finite (conditional) variances:

$$E[\epsilon_{iNt}^2] < \infty,$$

for all i, N and t. Because we will need to estimate consistently second moments of the projection errors, we assume, in addition, that:

Assumption 2

$$E[\epsilon_{iNt}^4] < \infty,$$

for all i, N and t.

Assumption 1 is a bit stricter than what we really need[9], but we retain it because it is standard in asset pricing, and, therefore, will facilitate power analysis (to be dealt with in the next section). On the other hand, Assumption 1 leaves the tail of the Nth sequence (*i.e.*, variables indexed $i = N+1,\ldots$) unrestricted when $N < \infty$. In other words, the regression errors beyond the Nth one could be arbitrary, no matter how large N. In particular, the largest eigenvalue of their variance–covariance matrix in the Nth sequence may be arbitrarily large, whereas it may be finite in the $N+1$st sequence, and again infinite in the $N+2$nd sequence. We want to avoid such erratic behavior by imposing that the variance–covariance matrix of *complete* adjacent sequences converge sufficiently fast as $N \rightarrow \infty$. In particular, we assume the following.

Assumption 3 *Let*

$$s_N = \sup_M\{\frac{\sup\{\lambda(\Sigma_{MN})\}}{M}\}.$$

[9] It is sufficient to assume that the eigenvalues diverge at a rate strictly less than N.

Then:

$$\lim_{N\to\infty} N s_N = 0.$$

Notice that Assumption 3 does not imply Assumption 1. As pointed out before, the latter is actually not necessary for our statistic to have the right asymptotic properties.

We assume that the projection errors are exchangeable.

Assumption 4 *For all N, the sequence of random vectors $(\epsilon_{1Nt},\ t = 1, ..., T)$,... $(\epsilon_{NNt},\ t = 1, ..., T)$, $(\epsilon_{N+1Nt},\ t = 1, ..., T)$,... is exchangeable.*

Let $i_1, i_2, ..., i_n$ be a collection of n positive integers. Exchangeability means that

$$\{(\epsilon_{i_1Nt}, t = 1, ..., T), (\epsilon_{i_2Nt}, t = 1, ..., T), ..., (\epsilon_{i_nNt}, t = 1, ..., T)\}$$

has the same joint (conditional) distribution as

$$\{(\epsilon_{i_{\pi 1}Nt}, t = 1, ..., T), (\epsilon_{i_{\pi 2}Nt}, t = 1, ..., T), ..., (\epsilon_{i_{\pi N}Nt}, t = 1, ..., T)\}$$

for all permutations π and each n ($n < \infty$). Assumption 4 imposes a cross–sectional homogeneity restrictions on the regression errors for a given N. It means that their distributional characteristics are indistinguishable *a priori*. Both cross–sectional and time dependence, however, are allowed.[10] Also, no time series homogeneity is imposed: the errors can be heteroscedastic over time, for instance.[11]

To test mean–variance efficiency, we use the following estimate of the intercept in the projection of asset i's excess return onto p_N's excess return:

$$\hat{\alpha}_{iN} = \frac{1}{T}\sum_{t=1}^{T}(r_{it} - \hat{\beta}_{iN} r_{p_N t}), \tag{3}$$

where $\hat{\beta}_{iN}$ is the least squares estimate of the slope in a regression *without intercept* of i's excess return onto that of p.[12] Using (1), we can relate $\hat{\alpha}_{iN}$ to the projection errors:

$$\hat{\alpha}_{iN} = \alpha_{iN}\Big(1 - \frac{1}{T}\frac{(\sum_{t=1}^{T} r_{p_N t})^2}{\sum_{t=1}^{T} r_{p_N t}^2}\Big) + \frac{1}{T}\sum_{t=1}^{T} w_{Nt}\epsilon_{iNt}, \tag{4}$$

[10] Assumption 4, combined with Assumption 1, implies that asset returns satisfy a *strict* factor structure asymptotically (*i.e.*, the regression errors become uncorrelated in cross–section as $N \to \infty$).

[11] As mentioned in the Introduction, we could also have obtained the same asymptotic properties for our test statistic by assuming that the regression errors are strong mixing. We opted not to go that route, because strong mixing is a time series concept (depends on a natural order) and the independence it imposes is too strong in view of the cross-sectional empirical properties of asset returns.

[12] If we had taken an estimate based on a regression *with intercept*, the expression for $\hat{\alpha}_{iN}$ would have been more involved, and the analysis correspondingly complex. This alternative approach would matter, even asymptotically. We plan to investigate this issue in future work.

where

$$w_{Nt} = 1 - r_{pNt} \frac{\sum_{\tau=1}^{T} r_{pN\tau}}{\sum_{\tau=1}^{T} r_{pN\tau}^2}.$$

Define:

$$\epsilon_{iN}^T = \frac{1}{T} \sum_{t=1}^{T} w_{Nt} \epsilon_{iNt}.$$

Obviously, $E[(\epsilon_{iN}^T)^2] < \infty$, $E[(\epsilon_{iN}^T)^4] < \infty$, and, because w_{Nt} does not depend on i, the sequences of random variables $\epsilon_{1N}^T,\ldots, \epsilon_{NN}^T, \epsilon_{N+1N}^T,\ldots$ are exchangeable.

The asymptotic properties of our statistic are based on an extension of De Finetti's theorem (Loève [1960, p. 365]) that exchangeable random variables are independently and identically distributed *conditional on* a (possibly infinite–dimensional) random variable. De Finetti's conditioning variable operates as a factor in factor analysis: it purges the data of cross–sectional dependence. The problem is that little is usually known about it.[13] The assumption that the regression errors become idiosyncratic as the size of the cross–section increases, however, restricts the conditioning variable such that the regression errors obey a central limit theorem *unconditionally.* Since the estimates of the intercepts are linearly related to the regression errors, the central limit theorem carries over to the former. Let us show this now.

Let $D(-\infty, +\infty)$ denote the set of distribution functions on R. Exchangeability implies:

Lemma 1 (De Finetti; Loève) *For all N, there exists a random variable Y_N taking values in $D(-\infty, +\infty)$ such that $\{\epsilon_{iN}^T\}_{i=1}^{\infty}$ are i.i.d. conditional on Y_N.*

(All proofs are in the Appendix.) Let $P_{MN}(d(\epsilon_{1N}^T, \ldots, \epsilon_{MN}^T, Y_N))$ be the joint probability measure induced by $(\epsilon_{1N}^T, \ldots, \epsilon_{MN}^T, Y_N)$ on $R^N \times D(-\infty, +\infty)$. We can always decompose this probability measure as follows:

$$P_{MN}(d(\epsilon_{1N}^T, \ldots, \epsilon_{MN}^T, Y_N)) = Q_{Y_N,M}(d(\epsilon_{1N}^T, \ldots, \epsilon_{MN}^T)) S_N(dY_N).$$

S_N is the marginal probability measure of (induced by) Y_N. $Q_{Y_N,M}$ is the joint conditional probability measure of (induced by) the first M ϵ_{iN}^Ts, where Y_N is the conditioning variable. Lemma 1 implies:

$$Q_{Y_N,M}(d(\epsilon_{1N}^T, \ldots, \epsilon_{MN}^T)) = \prod_{i=1}^{M} Q_{Y_N,1}(d\epsilon_{iN}^T).$$

[13]Absent restrictions on the conditioning variable, De Finetti's theorem is of interest mainly to Bayesian statisticians (see, *e.g.*, Florens, Mouchart and Rolin [1990, Section 9.3.2]).

The conditioning variable in Lemma 1 can be shown to be the distribution function of ϵ_{iN}^T (any i) corresponding to $Q_{Y_N,1}$.[14] Y_N can be consistently estimated from the empirical distribution function. This means that, for rational ν:

$$Y_N(\nu) = \lim_{M\to\infty} \frac{1}{M}\sum_{i=1}^{M} 1_{\{\epsilon_{iN}^T \leq \nu\}},$$

(almost surely). Using Riemann–Stieltjes integration, define

$$\bar{Y}_N = \int_R \nu dY_N(\nu).$$

$\bar{Y}_N$ is the conditional expectation of ϵ_{iN}^T (any i) with respect to Y_N, *i.e.*, $E[\epsilon_{iN}^T|Y_N]$. The sample average $\sum_{i=1}^{M} \epsilon_{iN}^T/M$ provides a strongly consistent estimator of $\bar{Y}_N$.

The conditional independence immediately implies a central limit theorem. In particular, let

$$V_N^* = \frac{1}{N}\sum_{i=1}^{N}(\epsilon_{iN}^T - \frac{1}{N}\sum_{j=1}^{N}\epsilon_{jN}^T)^2. \tag{5}$$

Conditional on $\{Y_N\}_{N=1}^{\infty}$, the statistic

$$z_N^{**} = \frac{1}{\sqrt{N}}\frac{1}{\sqrt{V_N^*}}\sum_{i=1}^{N}(\epsilon_{iN}^T - \bar{Y}_N) \tag{6}$$

converges weakly to a standard normal distribution as $N \to \infty$. Letting z denote a standard normal random variable, and adapting Billingsley [1986, p. 245], this means:

$$\lim_{N\to\infty} E[h(z_N^{**})|Y_N] = E[h(z)],$$

(almost surely), for all bounded, continuous functions h.

At first, this result may seem useful. Indeed, provided $\alpha_{iN} = 0$, all i and N, a central limit theorem for the cross–sectional average of the estimated intercepts (the $\hat{\alpha}_{iN}$s) would follow from the relation between the latter and the projection errors (the ϵ_{iN}^Ts; see Equation (4)). Consequently, it appears as if we obtained the result we sought for.

Unfortunately, this is not so, because the statistic z_N^{**} depends on $\bar{Y}_N$, which is unknown. The assumption that the regression errors become idiosyncratic as the size of the cross–section increases (Assumption 1), however, restricts $\bar{Y}_N$ in the limit. Since

[14]Hence, if, for instance, $\{\epsilon_{iN}^T\}_{i=1}^{\infty}$ is a sequence of *unconditionally i.i.d.* random variables, then Y_N is constant, *i.e.*, does not vary across samples, and equals the unconditional distribution of any element in the sequence.

Assumption 3 requires sequences to quickly behave similarly, the restriction will apply to $\bar{Y}_N$ for finite N as well (provided N is sufficiently large). The restriction is such that the distribution of z_N^{**} becomes independent of $\bar{Y}_N$ for large N. Hence, we obtain an *unconditional* central limit theorem. The nature of the restriction on $\bar{Y}_N$ is as follows.

Lemma 2 *For all $\delta > 0$,*

$$\lim_{N\to\infty} S_N\{\sqrt{N}|\bar{Y}_N| > \delta\} = 0.$$

Consequently, while Y_N may vary across samples, $\bar{Y}_N$ converges faster than $\sqrt{N}$ to zero (in probability).

Redefining the statistic as follows (V_N^* is as defined in (8)),

$$z_N^* = \frac{1}{\sqrt{N}}\frac{1}{\sqrt{V_N^*}}\sum_{i=1}^{N} \epsilon_{iN}^T, \tag{7}$$

we obtain an unconditional central limit theorem. Remembering that z denotes a standard normal random variable, and letting '$\Rightarrow$' denote 'converges weakly to',

Theorem 1

$$z_N^* \Rightarrow z.$$

The desired result follows from the relation between the estimated intercepts and the projection errors (Equation (4)). Define[15]:

$$V_N = \frac{1}{N}\sum_{i=1}^{N}(\hat{\alpha}_{iN} - \frac{1}{N}\sum_{j=1}^{N}\hat{\alpha}_{jN})^2 \tag{8}$$

and

$$z_N = \frac{1}{\sqrt{N}}\frac{1}{\sqrt{V_N}}\sum_{i=1}^{N}\hat{\alpha}_{iN}. \tag{9}$$

Corollary 1 *Under the null hypothesis (Equation (2)),*

$$z_N \Rightarrow z.$$

[15]Alternatively, we could have defined V_N as $\frac{1}{N}\sum_{i=1}^{N}\hat{\alpha}_{iN}^2$. Because of Lemma 2, the results would not change. In the empirical application, we used this alternative definition.

This corollary states that one can use the cross–sectional average of the estimated intercepts in excess return projections to test mean-variance efficiency. The size of the cross–section is exploited. The length of the time series, however, is kept fixed, and could be as small as two (2) observations. Notice also that no distributional assumptions (*e.g.*, normality) were imposed. While we assume that the errors in the regression of excess returns of individual assets onto that of the candidate portfolio become idiosyncratic as the size of the cross–section increases, we require neither cross–sectional nor time series independence. The statistical properties of the regression errors, however, are assumed to be indistinguishable. This is justified by the stringent limits that stationarity imposes on the time series length.

Our test statistic essentially uses the same inputs as the traditional portfolio–based procedure: the cross–sectional average of the estimated intercepts. The standard method directly investigates the distributional characteristics of this average, sometimes allowing the time series length to increase without bound (see, *e.g.*, Shanken [1992]). In contrast, the properties of our statistic have been derived after multiplying this average by the square root of the cross–sectional size. As explained in the next section, this difference generates the incremental power of our test.

3 Power

We now investigate the power of our testing procedure against the alternative that the limit portfolio p is not mean–variance efficient. We continue to assume that p is well–diversified and any risk uncorrelated with p is idiosyncratic. At the end of this section, we briefly discuss the implication of relaxing this assumption.

Under the alternative hypothesis, we want to allow for the possibility that there does not exist a well–defined mean–variance trade–off. While the assumption that a nontrivial mean–variance trade–off exists is innocuous in finite markets (Roll [1977]), it cannot be taken for granted in very large ('infinite') markets. Chamberlain and Rothschild [1983] prove that this requires the absence of arbitrage opportunities. Consequently, under the alternative hypothesis, we do not want to rule out the possibility that (potentially sophisticated) arbitrage opportunities exist. The power of our statistic hinges on whether or not there are such arbitrage opportunities. If there are, our test will have unit asymptotic power. If there are none, our test can still have power, but strictly less than one. We will investigate both cases.

One type of obvious arbitrage opportunity will not be allowed, however. Since we assume that risk uncorrelated with p is idiosyncratic, any other well–diversified portfolio with positive variance must be perfectly correlated with p. This restricts the expected return on the former. In particular, the intercept in a regression of the excess return of a well-diversified portfolio onto that of p must be zero. But this intercept is in fact a weighted average of the intercepts in analogous regressions of the excess return on the component assets. Consequently, we restrict weighted averages (using well-diversified weights) of the intercepts to be zero. Specifically:

Assumption 5 *Let x_{iN} ($i = 1, ..., N$, $N = 1, 2, 3, ...$) be sequences of real numbers such that (i) $\sum_{i=1}^{N} x_{iN} = 1$, all N, (ii) $\lim_{N\to\infty} \sum_{i=1}^{N} x_{iN}^2 = 0$. Then:*

$$\lim_{N\to\infty} \sum_{i=1}^{N} x_{iN}\alpha_{iN} = 0.$$

Standard tests of mean-variance efficiency regress the excess return on a portfolio of assets onto that of the candidate portfolio. The weights of the former are chosen to be small, so that the estimation error is minimal. If assets are priced such that there are no arbitrage opportunities between (almost) perfectly correlated well-diversified portfolios, the intercept of the regression of excess returns can be expected to be close to zero, independent of whether the regressor is mean-variance efficient. Consequently, it is not surprising that standard tests have very low power when a reasonably large number of assets are included in the regressand portfolio (see Gibbons, Ross and Shanken [1986]). In other words, under Assumption 5, the standard test procedure has effectively zero power. This is discussed in detail in Heston [1992]. In contrast, despite Assumption 5, our test statistic can have power. It may even have *unit* power.

Chamberlain [1983] has shown that the absence of any arbitrage opportunities is equivalent to the restriction that:

$$\lim_{N\to\infty} \sum_{i=1}^{N} \alpha_{iN}^2 < \infty \tag{10}$$

(this is Ross' Arbitrage Pricing Theory; see Ross [1976]). Because we rule out only the most obvious type of arbitrage opportunities, the sum in (10) may not be bounded. We will discuss how this affects the power of our test shortly.

Under the alternative hypothesis, the scaling factor V_N in the definition of z_N depends, among other things, on $\sum_{i=1}^{N} \alpha_{iN}^2/N$. In order to avoid pathological cases, we will restrict this average.

Assumption 6

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{iN}^2 < \infty$$

We are now ready to state the main result. Inspection of the definition of z_N will reveal that its asymptotic behavior under the alternative hypothesis (when some α_{iN} are nonzero for all N) depends crucially on the scaled average α_{iN}s:

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN}.$$

Indeed:

Theorem 2 *For all $B > 0$,*

$$\lim_{N \to \infty} P_{NN}\{|z_N| > B\} = 1$$

if and only if

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN}$$

diverges.

In words, our testing procedure has unit asymptotic power if the indicated condition is satisfied. Since

$$\lim_{N \to \infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{iN} = 0,$$

by Assumption 5, the condition for unit asymptotic power can be interpreted as requiring that the average value of the intercept converge at a rate slower than the square root of the size of the cross–section.

The condition for unit power is, however, violated when no arbitrage opportunities are present.

Lemma 3 *If*

$$\lim_{N \to \infty} \sum_{i=1}^{N} \alpha_{iN}^2 < \infty,$$

then

$$\lim_{N \to \infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN} < \infty.$$

Consequently, our test statistic does not have unit asymptotic power against the alternative that the candidate portfolio is not mean–variance efficient and there are no arbitrage opportunities, as stated in the following corollary.

Corollary 2 *If there are no arbitrage opportunities, i.e., if*

$$\lim_{N\to\infty} \sum_{i=1}^{N} \alpha_{iN}^2 < \infty,$$

then, for all $B > 0$,

$$\lim_{N\to\infty} P_{NN}\{|z_N| > B\} < 1.$$

Corollary 2 does not imply that our test has *zero* asymptotic power. This would be the case if

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN} = 0. \tag{11}$$

Absence of arbitrage opportunities does not necessarily imply (11), as the following example illustrates.

Example: Let $\alpha_{iN} = (iN)^{-\frac{1}{4}}$. Then:

$$\lim_{N\to\infty} \sum_{i=1}^{N} \alpha_{iN}^2 = \int_0^1 \frac{1}{\sqrt{\nu}} d\nu = 2 < \infty,$$

and

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN} = \int_0^1 \frac{1}{\nu^{\frac{1}{4}}} d\nu = \frac{4}{3} > 0.$$

Incidentally, this configuration satisfies Assumption 5:

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{iN} = 0.$$

□

We cannot exclude, however, the possibility that our test statistic has zero power. Nevertheless, it constitutes an improvement over standard tests, which, as mentioned before, should be expected to have zero power because of Assumption 5. The discussion of the power of our test should have revealed why it dominates: it is effectively a test of whether

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN} = 0.$$

In contrast, the standard procedure tests whether

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{iN} = 0.$$

Even if the latter holds, the former may not.

Actually, the fact that our test may have zero power in certain cases can be exploited to confirm the empirical validity of the theoretical distributional properties of our statistic. This is a unique opportunity to obtain validation of theoretical statistical results on the basis of the data itself instead of Monte Carlo analysis. In Section 5, we will test the mean–variance efficiency of an equally weighted index of all NYSE and AMEX securities. The test will be run on virtually the complete cross–section with which the index is constructed. But then the cross-sectional average of the intercepts ($\sum_{i=1}^{N} \alpha_{iN}/N$) should be close to zero no matter what the size of the cross–section (N) is. Consequently,

$$\frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN} \cong 0,$$

all N. Because of this, our test has essentially zero power, and our statistic should be distributed as if the null hypothesis were true. In particular, it ought to be standard normally distributed, provided the remaining assumptions still hold and the cross–sectional size is large enough to rely on asymptotic results. Section 5 will discuss whether this can be affirmed.

As Chamberlain and Rothschild [1983] showed, there will not exist a well–defined mean–variance trade–off if arbitrage opportunities are present. It would be interesting to have an indication of what causes a rejection of mean–variance efficiency: is it merely because the candidate portfolio is not mean–variance efficient, or does it follow from the fact that there are arbitrage opportunities in the cross–section, and, hence, there does not exist a mean–variance efficient frontier? Because the distinction between these two cases hinges on whether

$$\lim_{N \to \infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN}$$

is finite or not, the law of the iterated logarithm could be helpful here. Consider the following scaling of z_N, denoted z_N^{lil}:

$$z_N^{lil} = \frac{1}{\sqrt{2 \log \log \sqrt{N}}} z_N. \tag{12}$$

Under our null hypothesis, the law of the iterated logarithm (Chung [1974], Section 7.5) implies:

$$\overline{\lim}_{N \to \infty} z_N^{lil} = 1$$

(almost surely). Under the alternative hypothesis, and assuming absence of arbitrage opportunities (which implies that $\lim_{N \to \infty} \sum_{i=1}^{N} \alpha_{iN}/\sqrt{N} < \infty$),

$$\overline{\lim}_{N \to \infty} z_N^{lil} = 1$$

(almost surely) as well. Consequently, we can decide that the presence of arbitrage opportunities, and, hence, the nonexistence of a mean–variance efficient frontier, caused a rejection if we notice that:

$$|z_N^{lil}| > 1.$$

Finally, let us comment on what happens if the candidate portfolio is not well-diversified under the alternative. In that case, the asymptotic behavior of z_N will depend on $\bar{Y}_\infty$. The latter changes across samples. Whenever $\bar{Y}_\infty$ is nonzero, z_N will diverge. Since $\bar{Y}_\infty$ will be nonzero with positive probability, our test will be powerful.

4 Application

We implement our testing procedure on the CRSP dataset of monthly returns on NYSE and AMEX common stocks, over periods of five years. This should shed light on three questions: (i) Does our procedure generate the incremental power over the standard portfolio-based test that the theoretical analysis claims it has? (ii) Do rejections indicate that there are arbitrage opportunities, *i.e.*, that Ross' Arbitrage Pricing Theory is violated? (iii) What is the effect of excluding thinly traded, newly listed and delisted securities? In this section, we will answer questions (i) and (ii). Question (iii) will be dealt with in Section 6.

We test the mean-variance efficiency of the CRSP value weighted index. Value weighting does not necessarily lead to a well-diversified portfolio as the size of the cross-section increases; we merely assume that it will. Table 1 provides, for each subperiod, descriptive statistics of the returns on the candidate portfolio, as well as the returns on the equally weighted CRSP index (to be used in Section 5) and the riskfree rate (the one-month Treasury bill yield, from the Fama-Bliss files in the CRSP dataset).

We will directly compare the results from our testing procedure to those from a multivariate test based on equally weighted size-ranked portfolios. The size-sorted portfolios are constructed using the rankings provided by CRSP. Table 2 displays summary statistics about the ten size-based portfolios (portfolio 1 covers the smallest decile; portfolio 10 the largest decile). The standard testing procedure is based on a F statistic (denoted W_u) constructed from the estimated intercepts in regressions of the excess returns on each size-ranked portfolio onto that of the candidate portfolio (see Gibbons, Ross and Shanken [1989]).

In implementing our own testing procedure, we generated several sets of results. The first set is based on a cross-section of only continuously-traded securities. In other words, the assets included in the test are only those that had no missing returns over the relevant five-year period. The second to fourth sets are based on several selections of all securities which traded at least ten of the sixty months. The fifth set covers all securities that traded at least ten of the sixty months. A comparison of the results across the five sets would reveal the extent of the bias that is introduced when excluding thinly traded, delisted and newly listed securities. For now, we will work with the fifth set (untruncated), delegating a discussion of the impact of selection bias to Section 6.

We also generated a parallel set of results where we required securities to have returns for only five (instead of ten) months in order to be included in the sample. The conclusions did not change, and, hence, we do not report them here (the interested reader can obtain the results by contacting us). As with the tests of mean–variance efficiency over the month of January (Table 4 below), this illustrates the power of our test: only five observations (in time) were needed. While we can limit the time series dimension to a minimum, we exploit the cross–sectional size at a maximum.

Let us analyze the results for the value weighted CRSP index. Table 3 lists, for each subperiod: (i) the traditional W_u test, based on the size–ranked portfolios, (ii) the number of securities in the CRSP dataset that have at least one return over the subperiod, (iii) the cross–sectional size of all securities that traded at least ten months (N), (iv) the corresponding z_N statistic (Corollary 1) and its p-value, (v) the corresponding z_N^{lil} statistic (Equation (12)).

Let us focus on some of the most striking aspects of the results.

1. The traditional W_u test is unable to reject mean–variance efficiency in many subperiods, confirming the findings in Gibbons, Ross and Shanken [1989].
2. Our test statistic is far more powerful than the standard one, rejecting in all but one subperiod.
3. The z_N^{lil} statistic is often above one (in absolute value), indicating that the sum of squared intercepts does not satisfy Ross' Arbitrage Pricing Theory restriction, and, hence, that there may be arbitrage opportunities.

The first point should be put into perspective: it obtains despite our including AMEX securities in the construction of the size–ranked portfolios. Gibbons, Ross and Shanken [1989] only used NYSE securities. The typical size of an AMEX–listed company is much smaller than that of a NYSE–listed company, hence, we would have expected to obtain much more evidence of a "size effect" (rejection of the mean–variance efficiency of the value–weighted portfolio based on size–ranked portfolios). But we did not.

The second point proves that our statistic generates the additional power that a theoretical analysis (Section 3) claims it has. It may still be the case that

$$\lim_{N\to\infty} \frac{1}{N} \sum_{i=1}^{N} \alpha_{iN} = 0$$

(which is what the traditional test verifies), but our statistic rejects that

$$\lim_{N\to\infty} \frac{1}{\sqrt{N}} \sum_{i=1}^{N} \alpha_{iN} = 0,$$

thereby rejecting mean–variance efficiency.

The third point must be qualified. If z_N equals 2, N must exceed 2,618,500 for z_N^{lil} to be below 1. Because of this, given our sample sizes, a rejection of mean–variance efficiency would automatically imply a rejection of Ross' Arbitrage Pricing Theory as well. Consequently, we cannot be very confident in the conclusions we draw from the z_N^{lil} statistic.

We also investigated the mean–variance efficiency of the CRSP value weighted portfolio over the month of January. Table 4 displays the results for our procedure (z_N and z_N^{lil} statistics). Results for the traditional test statistic (W_u) are provided only for the complete sample period (because subperiods cover only five observations, too few to calculate W_u). Mean–variance efficiency of the value weighted CRSP index can be rejected for all subperiods. The intensity of the rejections seems to indicate that there are arbitrage opportunities (z_N^{lil} is always above one).

Notice the positive sign of z_N: it means that the average estimated intercept is positive, indicating that the average security is located above the "securities market line" defined by the value weighted CRSP index. The standard test, based on size–sorted portfolios, rejects for the complete period. The average estimated intercept of the ten size–based portfolios (not reported) is positive as well.

The rejections over January are noteworthy because of the extremely short time series length (five observations!) that is used to estimate the individual intercepts. It attests that our procedure is very powerful indeed.

One additional remark should be made. If the "market portfolio" is the limit of the CRSP value weighted portfolio as the size of the cross–section of securities increases, the results in Tables 3 and 4 can be interpreted as rejecting (the Sharpe Lintner version of) the Capital Asset Pricing Model.

5 Validating the Distributional Properties

As mentioned when we discussed power (Section 3), we can verify the asymptotic properties of our statistic by choosing the candidate portfolio such that our test has zero power. In particular, we can take the candidate portfolio to be the equally weighted portfolio of roughly the same securities that are used in the test.

Table 5 reports the results of applying our test on the equally weighted CRSP index in the cross–section of all securities that traded at least ten months. While the equally weighted CRSP index contains more securities, a comparison of columns 2 (which lists the number of securities in the CRSP dataset that had at least one return over the relevant subperiod) and 3 (which lists N, the size of the cross-section used in the tests) reveals that the difference is minimal. Consequently, the power of our test should be close to zero.

Looking at Table 5, we reject at the 5% level in only one of the thirteen subperiods. This confirms the theoretical analysis: if z_N is indeed standard normal, the probability of observing one or more rejections in thirteen random draws is 35%. The (time series) sample average equals 0.26, with a standard error of 0.28 (hence, it is insignificantly different from zero). The sample standard deviation equals 0.99. The corresponding χ^2 statistic has a value of 12.74, which, at twelve degrees of freedom, is close to the expected value of this statistic under the hypothesis that the population standard deviation equals one. The sum of the squared z_Ns should be χ^2 distributed with thirteen degrees of freedom. This sum equals 12.65, which again is close to its expected value under the null that z_N is standard normal. The skewness and kurtosis of the empirical distribution of z_N equal 0.21 and 2.31, respectively. The Kolmogorov–Smirnov statistic that tests standard normality equals 0.29, which corresponds to a p-level of roughly 15%. The Shapiro Wilk variance test of normality equals 0.94, with a p-level of about 40%.

There is, however, slight evidence against the hypothesis that z_N has a distribution that is symmetric around zero (like the standard normal distribution). z_N is positive in ten out of thirteen subperiods. The probability of observing ten or more positive outcomes in thirteen random draws from a distribution that is symmetric around zero is 3%. The evidence of asymmetry was not present in the parallel set of results, based on securities that traded at least five (instead of ten) months.

6 Selection Bias

Our procedure provides an opportunity to investigate the impact of selection bias. Recent theoretical work (Brown, Goetzmann, Ibbotson and Ross [1992]) has indicated that this may have have been underestimated. Our test does not require continuously traded securities. In fact, we need only a minimum of two (time series) observations for each security in the cross-section.[16] In our empirical investigation, we tested the mean-variance efficiency of the equally weighted and value weighted portfolios on five samples:

1. all continuously traded stocks,
2. all continuously traded stocks, plus stocks that traded at least ten months and the first month (essentially the delisted securities),

[16] In Section 2, the theoretical properties of our statistic were derived under a fixed time-series length (T). It should be obvious that the results would still obtain if we allowed the time-series length to vary over the cross-section, subscripting T: T_i. The ensuing complexity kept us from writing Section 2 at such a level of generality.

3. all continuously traded stocks, plus stocks that traded at least ten months and the last month (essentially the newly listed securities),
4. all securities under 1, 2 and 3,
5. all securities that traded at least ten months (these include thinly traded securities as well).

Table 6 displays sample sizes and z_N statistics for each case. Undoubtedly, the effect of selection bias can be dramatic. For instance, if we had tested the mean–variance efficiency of the value-weighted CRSP index over the period 1971-75 on a subsample of continuously–traded stocks, we would have incorrectly inferred that it was inefficient. Since the signs of the average estimated intercept and z_N are the same, the results reveal that the average estimated intercept was negative for each group of excluded securities (delisted, newly listed and thinly traded issues). In fact, across subperiods, delisted securities more often *raise* the z_N statistic, whereas newly listed and thinly traded securities *lower* it. The positive effect of delisting on the average estimated intercepts must be related to takeovers. The negative effect of new listings confirms the finding in Ritter [1991].[17]

Does the traditional test based on size–sorted portfolios mitigate selection bias? At first, one would tend to answer this question affirmatively, because the return on the size–ranked portfolios is constructed from the returns on *all* securities that traded during a given month. This number, however, changes randomly from month to month, and, hence, it is not clear whether the average excess return on partially traded securities in the size–sorted portfolios will be picked up in the estimate of the intercept or, instead, absorbed by the noise term. An investigation of Tables 3 and 6 reveals that the traditional test is not always capable of recognizing selection bias. In the 1971-1975 period, the traditional test rejects mean–variance efficiency of the value weighted portfolio (Table 3), as does our test on a truncated sample of continuously traded securities (Table 6). Yet, our test on the complete sample (Tables 3 and 6) fails to reject.

As a check on the calculations, Table 7 compares the averages of the estimated intercepts for our test against those for the traditional test. The average estimated intercept of the ten (equally weighted) size–based portfolios (used to calculate W_u) should be close to that with which z_N is computed. Table 7 confirms this for both the value weighted and equally weighted CRSP index.

[17]One cannot necessarily conclude that delisted securities "outperform" the CRSP value weighted index, whereas newly listed and thinly traded issues "underperform", because the failure to reject mean–variance efficiency implies that the *true* (*i.e., ex ante*) intercepts must all be zero. To the extent that it can be identified *ex ante*, however, we may be able to reject mean–variance efficiency of the value weighted CRSP index on the subsample of newly issued securities. Notice that this would not imply that new issues are "overpriced". To determine the latter, we need an asset pricing model that has been validated empirically over the same subperiod.

7 Conclusion

This paper has introduced a new test of mean–variance efficiency. It is simple, yet powerful. It exploits the cross–sectional size of typical financial datasets, fixing the time dimension. We reject mean–variance efficiency of the value weighted portfolio on NYSE and AMEX stock returns. The test results are sensitive to deletion of stock that did not trade continuously, confirming recent fears that selection bias may have been underestimated.

One may argue that our rejections do not add to our empirical knowledge in view of the rejections using standard portfolio–based techniques and a test that nests mean–variance efficiency models in the presence of a riskfree asset within the more larger set of models of mean–variance efficiency without a riskfree asset (see MacKinlay [1987]). There are two objections, however. First, rejections in the latter context remain subject to the criticism that the portfolio formation criterion is arbitrary. Second, such tests are based on asymptotic results that require the time series length to increase without bound, because they verify *nonlinear* restrictions. It is generally agreed on, however, that the time series length should be limited because of suspicion that nonstationarities will confound the results.

It would be interesting, however, to investigate whether an equivalent of the traditional test of mean–variance efficieny in the absence of a riskfree asset could be derived in our context of exchangeable idiosyncratic risk. One cannot be too optimistic, because our appeal to De Finetti's theorem to generate an *unconditional* central limit theorem is highly specific. In particular, the asymptotic properties depend critically on the linear dependence of the test statistic on the idiosyncratic risk.

We have assumed that the errors from regressing the returns of an individual asset onto those of the candidate portfolio are exchangeable in cross–section. Some may object to the homogeneity this implies, but we argued that the length of the time series does not allow us to model heterogeneity in much detail (in one case, we could use only five observations!), and, that, hence, individual idiosyncratic risk must be considered to be (not necessarily independent) draws from a homogeneous population, *i.e.*, exchangeable random variables. Empirical results endorse the statistical reasoning: when we tested the mean–variance efficiency in a case where we knew our procedure had zero power, we could not reject the hypothesis that our statistic is standard normally distributed.

One should view our imposing homogeneity on the idiosyncratic risk as a constructive assumption: it allows us to learn about the nature of the mean–variance efficient set despite the nonstationarities in the data. In this sense, the paper answers the following question: *is there any information about the mean–variance frontier in a cross–section of very short return records?*

For efficiency reasons, however, we may insist on modeling heterogeneity to a limited extent. Theory sometimes identifies variables that allow us to do so. The model of Bossaerts and Green [1988], for instance, suggests that size should be directly related to cross-sectional heterogeneity of idiosyncratic risk. This possibility could be worth further investigation.

We have essentially assumed an (approximate) one-factor model. Our analysis can easily be extended to incorporate models with multiple factors. This would amount to a test of whether a *combination* of portfolios is on the mean-variance efficient frontier. Gibbons, Ross and Shanken [1989], for instance, discuss this possibility in the context of traditional tests. Given the rejection of mean-variance efficiency of a single portfolio we discovered in the data, a test of whether a combination of portfolios is on the frontier should be the next empirical step.

Our tests verify unconditional mean-variance efficiency. There is both theoretical work (Hansen and Richard [1987]) and empirical work (*e.g.*, Bossaerts and Green [1988], Harvey [1991]) on testing *conditional* mean-variance efficiency. It would at first seem to be a simple logical extension to allow our statistic to test conditional mean-variance efficiency as well. Future research should indicate whether this is so.

Appendix

Proof of Lemma 1: See Loève [1960, p.365]. □

Proof of Lemma 2: Because Y_N is the almost sure limit of the sample average of the regression errors (for fixed N), we obtain the following upper bound on the probability that $\sqrt{N}|\bar{Y}_N| > \delta$:

$$
\begin{aligned}
& S_N\{\sqrt{N}|\bar{Y}_N| > \delta\} \\
= & \lim_{M\to\infty} P_{MN}\{\sqrt{N}|\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iN}^{T}| > \delta\} \\
\leq & \lim_{M\to\infty} \frac{NE|\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iN}^{T}|^2}{\delta^2}
\end{aligned} \tag{13}
$$

(by Chebychev's inequality). The expectation in the numerator of (13) can be bounded, as follows.

$$
\begin{aligned}
& E|\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iN}^{T}|^2 \\
= & E|\frac{1}{M}\frac{1}{T}\sum_{i=1}^{M}\sum_{t=1}^{T} w_{Nt}\epsilon_{iNt}|^2 \\
= & E|\frac{1}{T}\sum_{t=1}^{T} w_{Nt}\left(\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iNt}\right)|^2 \\
\leq & \frac{1}{T^2}\sum_{t=1}^{T}|w_{Nt}|^2 E|\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iNt}|^2 \\
& + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{t'>t}^{T}|w_{Nt}||w_{Nt'}|E|(\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iNt})(\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iNt'})| \\
\leq & \frac{1}{M}\sup\{\lambda(\Sigma_{MN})\}\left(\frac{1}{T^2}\sum_{t=1}^{T}|w_{Nt}|^2\right) \\
& + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{t'>t}^{T}|w_{Nt}||w_{Nt'}|\left(E|\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iNt}|^2 E|\frac{1}{M}\sum_{i=1}^{M}\epsilon_{iNt'}|^2\right)^{\frac{1}{2}} \\
\leq & \frac{\sup\{\lambda(\Sigma_{MN})\}}{M}\left(\frac{1}{T^2}\sum_{t=1}^{T}|w_{Nt}|^2\right) \\
& + \frac{\sup\{\lambda(\Sigma_{MN})\}}{M}\left(\frac{2}{T^2}\sum_{t=1}^{T}\sum_{t'>t}^{T}|w_{Nt}||w_{Nt'}|\right) \\
\leq & s_N\left(\frac{1}{T^2}\sum_{t=1}^{T}|w_{Nt}|^2 + \frac{2}{T^2}\sum_{t=1}^{T}\sum_{t'>t}^{T}|w_{Nt}||w_{Nt'}|\right)
\end{aligned}
$$

(the second and third inequalities follow from Chamberlain [1983, p.1321]; s_N in the last expression is defined in Assumption 3). Combining this result with Equation (13), letting $N \to \infty$ and appealing to Assumption 3 proves the Lemma. □

Proof of Theorem 1: Weak convergence follows if we can show (Billingsley [1986, p. 245]):

$$\lim_{N \to \infty} E[h(z_N^*)] = E[h(z)],$$

for all bounded, uniformly continuous functions h. Let $H < \infty$ be such that $|h| \leq H$. Consider:

$$\begin{aligned}
&|E[h(z_N^*) - h(z)]| \\
&\leq \ |E[h(z_N^{**}) - h(z)]| + E|h(z_N^{**}) - h(z_N^*)| \\
&\leq \ |E[E[h(z_N^{**})|Y_N] - h(z)]| + \varepsilon_N P_{NN}\{\sqrt{N}|\bar{Y}_N| \leq \delta_N \sqrt{V_N^*}\} \\
&\quad + 2HP_{NN}\{\sqrt{N}|\bar{Y}_N| > \delta_N \sqrt{V_N^*}\}
\end{aligned}$$

(the second inequality follows from the uniform continuity and boundedness of h). As $N \to \infty$, the first term in the last expression converges to zero (because of conditional weak convergence of z_N^{**} to z). Set: $\varepsilon_N = 1/N$. Then, by Lemma 2 and the boundedness of $\sqrt{V_N^*}$ (in P_{NN}), the second and third terms converge to zero as well. □

Proof of Corollary 1: Follows immediately from (i) substitution of (4) into the formula of z_N to obtain z_N^*, noting that $\alpha_{iN} = 0$ (all i), and (ii) $z_N^* \Rightarrow z$ (Theorem 1). □

Proof of Theorem 2: We first check whether V_N remains bounded (in P_{NN}). Let

$$c = 1 - \frac{1}{T}\frac{(\sum_{t=1}^T r_{p_N t})^2}{\sum_{t=1}^T r_{p_N t}^2}.$$

Then:

$$\begin{aligned}
V_N &= \frac{1}{N}\sum_{i=1}^N \hat{\alpha}_{iN}^2 - \left(\frac{1}{N}\sum_{i=1}^N \hat{\alpha}_{iN}\right)^2 \\
&= c^2 \frac{1}{N}\sum_{i=1}^N \alpha_{iN}^2 + 2c\frac{1}{N}\sum_{i=1}^N \alpha_{iN}\epsilon_{iN}^T + \frac{1}{N}\sum_{i=1}^N (\epsilon_{iN}^T)^2 \\
&\quad - \left(\frac{1}{N}\sum_{i=1}^N \hat{\alpha}_{iN}\right)^2 .
\end{aligned} \tag{14}$$

By Assumption 6, the first term in (14) remains finite. Using the same arguments as in the proof of Corollary 1, the last term converges to zero (in probability). As to the second term in (14), define weights x_{iN}, as follows:

$$x_{iN} = \frac{1}{N} + \frac{\epsilon_{iN}^T}{N}.$$

Asymptotically, these become the weights of a well-diversified portfolio. To see this, note that, using the arguments of the proof of Lemma 2 (but appealing to Assumption 1 instead of Assumption 2),

$$\sum_{i=1}^{N} x_{iN} \rightarrow 1,$$

(in P_{NN}), and

$$\sum_{i=1}^{N} x_{iN}^2 = \frac{1}{N} + 2\frac{1}{N^2}\sum_{i=1}^{N} \epsilon_{iN}^T + \frac{1}{N^2}\sum_{i=1}^{N}(\epsilon_{iN}^T)^2 \rightarrow 0,$$

(in P_{NN}). Rewriting the second term of (14) using the x_{iN}s gives:

$$2c\frac{1}{N}\sum_{i=1}^{N} \alpha_{iN}\epsilon_{iN}^T = 2c\sum_{i=1}^{N} x_{iN}\alpha_{iN} - 2c\frac{1}{N}\sum_{i=1}^{N}\alpha_{iN}.$$

The last expression converges to zero, by Assumption 5. Finally, the third term in (14),

$$\frac{1}{N}\sum_{i=1}^{N}(\epsilon_{iN}^T)^2$$

converges (almost surely) to the (conditional) variance of ϵ_{iN}^T, by exchangeability.

Now, consider z_N.

$$\begin{aligned} z_N &= \frac{1}{\sqrt{N}}\frac{1}{\sqrt{V_N}}\sum_{i=1}^{N}\hat{\alpha}_{iN} \\ &= c\frac{1}{\sqrt{N}}\frac{1}{\sqrt{V_N}}\sum_{i=1}^{N}\alpha_{iN} + \frac{1}{\sqrt{N}}\frac{1}{\sqrt{V_N}}\sum_{i=1}^{N}\epsilon_{iN}^T. \end{aligned}$$

This expression diverges with unit probability if and only if

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\alpha_{iN}$$

diverges. □

Proof of Lemma 3: Without loss of generality, assume that $a_{iN} \geq 0$, all i and N. Let $[c]$ denote the integral part of a real number c. Write the expressions in the lemma as integrals:

$$\sum_{i=1}^{N}\alpha_{iN}^2 = \int_0^1 \alpha_{[\nu N]+1,N}^2 N d\nu,$$

$$\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\alpha_{iN} = \int_0^1 \alpha_{[\nu N]+1,N}\sqrt{N} d\nu.$$

Let A_N be the subset of the unit interval where $\alpha_{[\nu N]+1,N}$ (as a function of ν) is strictly larger than one. Then:

$$\begin{aligned}
&\frac{1}{\sqrt{N}}\sum_{i=1}^{N}\alpha_{iN} \\
\le\ & \int_{[0,1]\setminus A_N}\alpha_{[\nu N]+1,N}\sqrt{N}d\nu + \int_{A_N}\alpha^2_{[\nu N]+1,N}Nd\nu \\
\le\ & \int_{[0,1]\setminus A_N}\alpha_{[\nu N]+1,N}\sqrt{N}d\nu + \int_0^1\alpha^2_{[\nu N]+1,N}Nd\nu.
\end{aligned}$$

The first term in the last expression remains finite. The second term remains finite as well, by assumption. Hence, the whole expression converges. □

Proof of Corollary 2: Follows immediately from the proof of Theorem 2 and Lemma 3. □

References

Billingsley, P. (1986): *Probability and Measure.* New York: Wiley.

Black, F., M. Jensen and M. Scholes (1972): "The Capital Asset Pricing Model: Some Empirical Tests," in: M. Jensen, ed., *Studies in the Theory of Capital Markets.* New York: Praeger.

Bossaerts, P. and R. Green (1988): "A General Equilibrium Model of Changing Risk Premia: Theory and Tests," *Review of Financial Studies*, 2, 467-495.

Brown, S., W. Goetzmann, R. Ibbotson and S. Ross (1992): "Survivorship Bias in Performance Studies," *Review of Financial Studies*, 5, 553-580.

Chamberlain, G. (1983): "Funds, Factors and Diversification in Arbitrage Pricing Models," *Econometrica*, 51, 1305-1323.

Chamberlain, G. and M. Rothschild (1983): "Arbitrage, Factor Structure, and Mean–Variance Analysis on Large Asset Markets," *Econometrica*, 51, 1281-1304.

Chung, K.L. (1974): *A Course in Probability Theory.* Orlando: Academic Press.

Connor, G. and R. Korajczyk (1992): "A Test for the Number of Factors in an Approximate Factor Model," Northwestern University, Kellogg Graduate School of Management, Finance Department.

Fama, E. and J. MacBeth (1973): "Risk, Return and Equilibrium: Empirical Tests," *Journal of Political Economy*, 71, 607-636.

Florens, J.P., M. Mouchart and J.-M. Rolin (1990): *Elements of Bayesian Statistics.* New York and Basel: Marcel Dekker.

Gibbons, M. (1982): "Multivariate Tests of Financial Models: A New Approach," *Journal of Financial Economics*, 10, 3-27.

Gibbons, M., S. Ross and J. Shanken (1989): "A Test of the Efficiency of a Given Portfolio," *Econometrica*, 57, 1121-1152.

Gibbons, M. and J. Shanken (1987): "Subperiod Aggregation and the Power of Multivariate Tests of Portfolio Efficiency," *Journal of Financial Economics*, 19, 389-394.

Hansen, L. and S. Richard (1987): "The Role of Conditioning Information in Deducing Testable Restrictions Implied by Dynamic Asset Pricing Models," *Econometrica*, 55, 559-586.

Harvey, C. (1991): "The World Price of Covariance Risk," *Journal of Finance*, 46, 111-157.

Heston, S. (1992): "Testing Approximate Linear Asset Pricing Models," Yale University, School of Organization and Management.

Lindley, D.V. and M.R. Novick (1981): "The Role of Exchangeability in Inference," *Annals of Statistics*, 9, 45-58.

Lo, A. and C. MacKinlay (1990): "Data–Snooping Biases in Tests of Financial Asset Pricing Models," *Review of Financial Studies*, 3, 431-468.

Loève, M. (1960): *Probability Theory.* Princeton, NJ: Van Nostrand.

Ritter, J. (1991): "The Long–Run Performance of Initial Public Offerings," *Journal of Finance*, 46, 3-28.

Roll, R. (1977): "A Critique of the Asset Pricing Theory's Tests. Part I: On Past and Potential Testability of the Theory," *Journal of Financial Economics*, 4, 129-176.

Ross, S. (1976): "The Arbitrage Theory of Capital Asset Pricing," *Journal of Economic Theory*, 13, 341-360.

Schwert, W. and P. Seguin (1990): "Heteroskedasticity in Stock Returns," *Journal of Finance*, 45, 1129-1156.

Shanken, J. (1992): "On the Estimation of Beta–Pricing Models," *Review of Financial Studies*, 5, 1-34.

Table 1
Summary Statistics: Candidate Portfolios and Risk Free Rate

	Value Weighted CRSP Index		Equally weighted CRSP Index		Risk Free Rate	
Subperiod	mean	s. dev.	mean	s. dev.	mean	s. dev.
1926-1930	0.55	6.02	-0.15	6.85	0.30	0.07
1931-1935	1.02	12.30	3.19	18.32	0.03	0.06
1936-1940	0.38	7.75	0.79	11.39	0.01	< 0.01
1941-1945	1.44	3.30	2.73	5.76	0.03	0.01
1946-1950	0.82	4.14	0.77	5.20	0.07	0.05
1951-1955	1.65	3.19	1.48	3.26	0.12	0.06
1956-1960	0.84	3.37	0.91	3.38	0.20	0.06
1961-1965	1.08	3.42	1.25	4.18	0.25	0.05
1966-1970	0.39	4.39	0.96	6.56	0.44	0.08
1971-1975	0.32	5.10	0.10	7.76	0.47	0.13
1976-1980	1.38	4.47	2.48	6.24	0.61	0.21
1981-1985	1.20	4.11	1.48	4.71	0.71	0.20
1986-1990	1.07	5.27	0.44	5.56	0.50	0.11

Remarks: Sample averages ("mean") and standard deviations ("s. dev.") of monthly returns are displayed. Estimates are multiplied by 100.

Table 2
Sample Average Monthly Returns of 10 Size-Ranked Portfolios
Raw Returns

Subperiod	s 1	s 2	s 3	s 4	s 5	s 6	s 7	s 8	s 9	s 10
1926-1930	-0.20	0.06	0.42	0.90	1.18	0.71	1.00	0.57	1.29	1.51
1931-1935	7.48	5.31	3.68	3.37	2.95	2.32	2.19	1.40	1.68	1.15
1936-1940	1.35	1.25	0.68	0.86	0.92	0.78	0.75	0.50	0.41	0.41
1941-1945	6.22	3.56	3.03	2.66	2.34	2.80	2.16	1.83	1.87	1.28
1946-1950	1.13	0.71	0.74	0.54	0.72	0.81	0.72	0.66	0.91	0.77
1951-1955	1.33	1.34	1.43	1.46	1.59	1.45	1.50	1.55	1.62	1.49
1956-1960	0.88	0.96	0.97	0.89	0.85	0.91	1.00	0.90	0.96	0.78
1961-1965	1.49	1.54	1.40	1.51	1.11	1.24	1.20	1.16	1.20	0.93
1966-1970	2.39	1.44	1.02	0.96	0.87	0.74	0.62	0.61	0.57	0.51
1971-1975	1.08	0.18	0.20	0.09	0.02	0.08	0.18	0.27	0.92	0.30
1976-1980	3.49	3.13	2.87	2.96	2.64	2.69	2.38	1.94	1.78	1.25
1981-1985	1.85	1.59	1.62	1.50	1.45	1.34	1.45	1.40	1.32	1.24
1986-1990	0.30	-0.44	-0.26	0.07	0.07	0.50	0.69	0.85	0.96	1.14

Remarks: s 1 = equally weighted portfolio of firms in the lowest (size) decile; s 10 = equally weighted portfolio of firms in the largest (size) decile. Returns are multiplied by 100.

Table 3
Tests of Mean-Variance Efficiency: Value Weighted CRSP Index

	Ten Size-Sorted Portfolios		All Stocks that Traded at Least Ten Months					
Subperiod	W_u	p-value	N_{max}	N	z_N	p-value	z_N^{lil}
1926-1930	3.84	<.01	842	783	-9.55	<.01	-6.16
1931-1935	0.98	.48	815	767	14.92	<.01	9.63
1936-1940	0.59	.82	859	828	5.06	<.01	3.25
1041-1945	1.08	.39	893	860	11.34	<.01	7.27
1946-1950	1.81	.08	1036	1005	-6.95	<.01	-4.41
1951-1955	1.39	.24	1119	1092	-2.47	.01	-1.56
1956-1960	0.86	.57	1231	1165	3.11	<.01	1.96
1961-1965	1.58	.14	2491	2317	3.66	<.01	1.29
1966-1970	2.38	.02	2975	2762	13.13	<.01	7.92
1971-1975	4.39	<.01	3059	2965	-1.71	.09	-1.03
1976-1980	2.02	.05	2975	2807	25.68	<.01	15.46
1981-1985	0.80	.63	2943	2692	5.80	<.01	3.50
1986-1990	3.20	<.01	3456	3158	-12.78	<.01	-7.65

Remarks: N_{max} = number of securities that had at least one return over the subperiod; N = size of the cross-section.

Table 4
Tests of Mean-Variance Efficiency: Value Weighted CRSP Index
(January Observations Only)

Subperiod	N	z_N	p-value	z_N^{lil}
1926-1930	449	5.06	<.01	3.39
1931-1935	655	9.70	<.01	6.32
1936-1940	659	9.42	<.01	6.14
1041-1945	762	13.40	<.01	8.65
1946-1950	839	12.71	<.01	8.16
1951-1955	976	8.99	<.01	5.72
1956-1960	969	18.93	<.01	12.05
1961-1965	1006	12.38	<.01	7.86
1966-1970	1708	24.75	<.01	15.27
1971-1975	2091	26.00	<.01	15.88
1976-1980	2047	25.91	<.01	15.84
1981-1985	1870	18.26	<.01	11.21
1986-1990	1644	3.51	<.01	2.17

Remarks: N = cross-sectional size; W_u, based on the returns of ten size-ranked portfolios over the complete sample period (1926-1990), equals 7.05 (significant at the 1% level).

Table 5
Empirical Distribution of z_N under the Null hypothesis

Subperiod	N_{max}	N	z_N	p-value
1926-1930	842	783	-0.646	0.52
1931-1935	815	767	0.168	0.87
1936-1940	859	828	0.132	0.90
1041-1945	893	860	0.616	0.54
1946-1950	1036	1005	0.240	0.81
1951-1955	1119	1092	0.017	0.99
1956-1960	1231	1165	0.441	0.66
1961-1965	2491	2317	-0.343	0.73
1966-1970	2975	2762	1.164	0.24
1971-1975	3059	2965	-1.800	0.07
1976-1980	2975	2807	2.470	0.01
1981-1985	2943	2692	0.061	0.95
1986-1990	3456	3158	0.857	0.39

Remarks: z_N statistics are displayed for tests of the mean-variance efficiency of the equally weighted CRSP index (the test has zero power against the alternative that this index is inefficient); the theoretical distribution of z_N is $N(0,1)$; N_{max} = number of securities that had at least one return over the subperiod; N = size of the cross-section (number of securities with return observations over at least ten months in the subperiod).

Table 6
The Impact of Sample Truncation on Tests of Mean-Variance Efficiency

		Subsample Size					Value Weighted CRSP Index					Equally Weighted CRSP Index				
Subperiod	N_{max}	N_1	N_2	N_3	N_4	N_5	z_{N1}	z_{N2}	z_{N3}	z_{N4}	z_{N5}	z_{N1}	z_{N2}	z_{N3}	z_{N4}	z_{N5}
1926-1930	842	381	437	649	705	783	-6.91	-6.22	-9.82	-9.39	-9.55	1.45	0.57	-0.31	-0.65	-0.65
1931-1935	815	597	667	638	708	767	16.24	14.20	16.55	14.51	14.92	-0.99	-0.48	-0.66	-0.26	0.17
1936-1940	859	634	687	756	809	828	4.12	3.82	5.40	5.00	5.06	-0.85	-0.91	0.32	0.10	0.13
1941-1945	893	751	781	821	851	860	13.38	10.95	13.71	11.31	11.34	-0.50	0.87	-0.64	0.71	0.61
1946-1950	1036	831	847	984	1000	1005	-5.97	-5.75	-7.10	-6.89	-6.95	0.93	1.06	0.14	0.29	0.24
1951-1955	1119	950	1006	1030	1086	1093	-2.79	-2.74	-2.72	-2.68	-2.47	-0.13	-0.07	-0.21	-0.15	0.02
1956-1960	1231	944	1028	1065	1149	1165	2.91	3.39	2.53	3.02	3.11	-0.44	0.18	-0.11	0.39	0.44
1961-1965	2491	946	1063	2019	2136	2317	4.41	5.01	3.00	3.47	3.66	3.73	4.84	-1.96	-1.10	-0.34
1966-1970	2975	1597	2034	2220	2657	2762	17.13	20.78	6.30	12.42	13.13	0.77	3.50	-1.63	0.73	1.16
1971-1975	3059	1959	2276	2527	2844	2965	6.29	3.09	1.41	-0.01	-1.71	5.35	3.32	-0.07	-0.48	-1.80
1976-1980	2975	1710	2172	2014	2476	2807	18.48	22.97	18.42	23.26	25.68	-4.82	1.07	-2.58	1.85	2.47
1981-1985	2943	1705	2197	2095	2587	2692	8.10	11.84	1.23	6.55	5.80	1.25	4.48	-2.86	0.89	0.06
1986-1990	3456	1542	2115	2430	3003	3158	-12.78	-7.20	-17.18	-12.79	-12.78	1.73	2.61	0.05	1.37	0.86

Remarks: z_N statistics are displayed for tests of mean-variance efficiency of the value weighted and equally weighted CRSP index; the tests are based on several subsamples, indicated with a subscript to the sample size (N); N_1 = size of cross-section of continuously traded securities; N_2 = N_1 + size of cross-section of securities that traded at least ten months but had a missing return for the last month ("delisted issues"); N_3 = N_1 + size of the cross-section of securities that traded at least ten months but had a missing return for the first month ("newly listed issues"); $N_4 = N_1 + (N_2 - N_1) + (N_3 - N_1)$; N_5 = size of cross-section of all securities that traded at least ten months; N_{max} = number of securities that had at least one return observation over the subperiod; z_N is standard normal, and, hence, statistically significant the 1% level if $|z_N| > 2.58$.

Table 7
Average Estimated Intercepts: Value Weighted and Equally Weighted CRSP Index

	Value Weighted CRSP Index		Equally Weighted CRSP Index	
Subperiod	Ten Size-Sorted Portfolios	All Stocks that Traded at least Ten Months	Ten Size-Sorted Portfolios	All Stocks that Traded at least Ten Months
1926-1930	-0.75	-0.80	-0.03	-0.05
1931-1935	1.77	1.77	-0.02	0.02
1936-1940	0.26	0.26	-0.00	0.01
1041-1945	0.78	0.72	-0.00	0.03
1946-1950	-0.21	-0.21	0.00	0.01
1951-1955	-0.08	-0.06	0.00	0.00
1956-1960	0.12	0.11	0.00	0.02
1961-1965	0.10	0.11	0.01	-0.01
1966-1970	0.58	0.30	-0.01	0.04
1971-1975	-0.01	-0.05	-0.03	-0.06
1976-1980	0.93	1.02	0.00	0.08
1981-1985	0.26	0.24	-0.01	0.00
1986-1990	-0.68	-0.59	-0.05	0.04

Remarks: Numbers shown are $\sum_{i=1}^{N} \hat{\alpha}_{iN}/N(\times 10^2)$, where N=cross-sectional size (N = 10 in the case of size-sorted portfolios).