

NBER TECHNICAL WORKING PAPER SERIES

KOLMOGOROV-SMIRNOV TESTS FOR DISTRIBUTION FUNCTION SIMILARITY
WITH APPLICATIONS TO PORTFOLIOS OF COMMON STOCK

Jack Meyer

Robert H. Rasche

Technical Working Paper No. 76

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 1989

We are indebted to Peter Schmidt for helpful comments and discussions. This research is part of NBER's research program in Financial Markets and Monetary Economics. Any opinions expressed are those of the authors not those of the National Bureau of Economic Research.

KOLMOGOROV-SMIRNOV TESTS FOR DISTRIBUTION FUNCTION SIMILARITY
WITH APPLICATIONS TO PORTFOLIOS OF COMMON STOCK

ABSTRACT

If the elements of the choice set in a decision model involving randomness are not arbitrary, but restricted appropriately, an expected utility ordering of them can be represented by a mean-standard deviation ranking function. These restrictions can apply to the form of, or can specify relationships among, the distribution functions. A particularly useful restriction is one which requires that elements in the choice set, when normalized to have a zero mean and unit variance, be identically distributed. No restriction is placed on the form of any individual distribution function.

This research empirically tests for this and other useful restrictions on the relationships among the elements of a set of random variables. Observations from the random variables are used to test whether or not they have distribution functions which are appropriately related to one another. The tests are applied to rate of return data for portfolios of common stock. The tests indicate that one cannot reject the hypothesis that the distribution functions of these portfolios are sufficiently similar to imply that the efficient set of portfolios for any risk averse expected utility maximizer is contained in the mean-standard deviation efficient set.

Jack Meyer
Economics Department
Michigan State University
East Lansing, MA 48824

Robert H. Rasche
Economics Department
Michigan State University
East Lansing, MI 48824

I. Introduction

In general discussions of expected utility (EU) decision models usually no assumptions are made concerning the form of, or the relationships among, the distribution functions describing the random alternatives. As a consequence, unless quadratic utility is assumed, the ordering under EU cannot be represented by one which depends on only the mean and standard deviation (MS) of the random alternatives.

On the other hand, if the elements in the choice set are not arbitrary, but restricted appropriately, an EU ordering can be represented by a MS ranking function. These restrictions can apply to the form of, or can specify relationships among, the distribution functions. Recently, Sinn [1983] and Meyer [1987] remind us of this, generalizing the normality restriction to one termed the "linear class" (Sinn) or the "location and scale" (Meyer) condition. This restriction requires that elements in the choice set, when normalized to have a zero mean and unit variance, be identically distributed. No restriction is placed on the form of any individual distribution function. Each of the random alternatives is equal in distribution to the others except for location and scale.

This research begins the process of empirically testing for this and other useful restrictions on the relationships among the elements of a set of random variables. Observations from the random variables are used to test whether or not they have distribution functions which are appropriately related to one another. This is done without restricting or specifying the form of the distribution functions involved. The tests are applied to rate of return data for portfolios of common stock.

The empirical analysis of stock portfolios is used to illustrate the test procedure, but the results that are obtained are also important in their own right. The tests indicate that one cannot reject the hypothesis that the distribution functions for rate of return on portfolios of common stock are sufficiently similar to one another to imply that the efficient set of portfolios for any risk averse expected utility investor is contained in the MS efficient set. This finding is not due to a special form for these distribution functions, but a result of their similarity. One implication of this finding is that the relatively simple MS portfolio building algorithms are appropriate for larger classes of investors than is indicated in the standard literature.

The paper is organized as follows. In the next section the approach used to model and test for relationships among random variables, and the specific hypotheses to be examined, are described. In section III, the Kolmogorov-Smirnov (KS) multi-sample test is suggested as an appropriate one when examining for these relationships among random rates of return on portfolios of common stock. Section IV presents the results from using this test to examine for the hypothesized similarities among these random alternatives using historical data. Finally, section V offers conclusions that can be drawn from the evidence presented here.

II. Model Structure

Random alternatives in many economic decision models involving randomness, result from an agent's selection of values for choice variables taking various random and nonrandom parameters as given. Thus, each random alternative has a common source of randomness, the

random parameters, and therefore is automatically related to the others. How they are related to one another depends on the structure of the equation which yields the random outcome as a function of the agent's choices and the random and nonrandom parameters. If this structure is appropriately restricted, then an EU ranking of the random alternatives can be represented by one depending only on their first two moments.

The general model structure dealt with here assumes that the random alternatives \bar{z}_i are equal in distribution to $a_i\bar{y} + b_i\bar{x} + c$, where a_i , b_i and c are nonrandom, and \bar{y} and \bar{x} are random. In this model, c , \bar{x} and \bar{y} are parameters outside the control of the decision maker, and a_i and b_i are selected by the decision maker from the feasible set. This structure implies that for given values for the parameters, the expected utility ranking of \bar{z}_i can be reduced to one which depends only on a_i and b_i . With further assumptions, these parameters can be made one to one with the mean and standard deviation of \bar{z}_i . This will be illustrated shortly.

While this structure may seem restrictive, it is sufficiently general to contain many of the economic models dealing with randomness found in the literature. In fact, most such models contain only one source of randomness, rather than the two allowed here.¹ Since the requirement is that \bar{z}_i be equal in distribution to $a_i\bar{y} + b_i\bar{x} + c$, the random variables \bar{y} and \bar{x} can differ across \bar{z}_i as long as the joint distribution function for \bar{y} and \bar{x} remains fixed. The two cases discussed next illustrate this possibility.

A special case of this general structure which follows the typical expected utility decision model with its one source of randomness,

assumes that \bar{z}_i is equal in distribution to $a_i + b_i\bar{x}$ where b_i is greater than zero. This is the structure pointed out by Sinn and Meyer in support of the linear class or location and scale condition since it implies that all \bar{z}_i are equal in distribution to one another except for location and scale. Sinn and Meyer show how EU and MS rankings of such random alternatives are related to one another. The specific formulation of this structure which we test for is:

Model 1: $\bar{z}_i = a_i + b_i\bar{x}_i$ where $b_i > 0$ and \bar{x}_i are identically distributed. Notice that the random term \bar{x}_i is allowed to differ across i , but its distribution function does not.

The second special case of interest allows two sources of randomness, but assumes that $\mu_y > 0$, $\mu_x = 0$ and that \bar{y} and \bar{x} are independent of one another. These restrictions do not imply the location and scale condition, but are sufficient to yield that the efficient set for any risk averse expected utility decision maker is contained in the MS efficient set.² The specific formulation of this structure which we test for is:

Model 2: $\bar{z}_i = a_i\bar{y}_i + b_i\bar{x}_i + c$ where $b_i > 0$, $\mu_y > 0$, $\mu_x = 0$, and the \bar{y}_i are identically distributed, the \bar{x}_i are identically distributed, and \bar{x}_i and \bar{y}_i are independent of one another. Again, the \bar{y}_i and \bar{x}_i can depend on i , but the assumptions imply that the joint distribution of (\bar{y}_i, \bar{x}_i) does not.

In summary, a general model structure which implies that the random alternatives the decision maker faces are related to one another in a two dimensional fashion has been identified. Two special cases of this structure which imply that the EU efficient set for all risk averse decision makers is contained in the MS efficient set are

also identified. The remaining sections of this paper discuss testing for such model structures, and whether or not rate of return data for portfolios of common stock can be adequately represented by a model of either of these two type.

III. Stock Portfolio Models and Kolmogorov-Smirnov Tests

Given $\{z_{it}\}$ a sample of observations from the random alternatives, the empirical question of concern is whether or not those observations are a likely result of Model 1 or 2 for any constants a_i , b_i , c , and random parameters \tilde{y}_i and \tilde{x}_i which satisfy the stated restrictions. How this question is addressed depends on the information available concerning the constants and random parameters.

If each of the nonrandom terms are known and the random parameters are observed concurrently with \tilde{z}_i , then each of the t observations must be checked for consistency with the specified model. This is the extreme case since all variables on the right and left side of the hypothesized model are observed.

A less extreme situation is one where the random parameters \tilde{y}_i and \tilde{x}_i are observed concurrently with \tilde{z}_i , but the a_i , b_i , and c are not known. For each i , it must be determined if $\{z_{it}\}$ lies in the three dimensional vector space spanned by $\{y_{it}\}$ and $\{x_{it}\}$ and a vector of ones, and if so, whether the coefficients satisfy the appropriate sign restrictions. Neither of these two situations involves statistical testing to confirm that the hypothesized model does or does not represent the process generating alternatives \tilde{z}_i .

A more interesting case is where one of the random parameters, say \tilde{x}_i , is not observed. In this situation the hypothesized model, with specified nonrandom terms and observations $\{z_{it}\}$ and $\{y_{it}\}$ can be

used to calculate values for $\{x_{it}\}$. The question is: would it be reasonable to observe these values as observations from an \bar{x}_i which satisfies the stated restrictions. Statistical tests can be used for this question. This is the situation which arises in analyzing the model representing the rate of return earned on a portfolio of common stock.

The single index version of the Capital Asset Pricing Model (CAPM) assumes a structure very similar to Models 1 and 2. It assumes that the rate of return for a portfolio i is given by $\bar{r}_i = \phi + \beta_i(\bar{r}_m - \phi) + \tilde{r}_i$, where ϕ is the risk free rate of return and \bar{r}_m is the random rate of return on the market portfolio. ϕ and \bar{r}_m are parameters to the agent. β_i is the covariance of the rate of return on the portfolio with the rate of return on the market portfolio divided by the market's variance, and is assumed to be constant. \tilde{r}_i represents nonsystematic risk associated with the portfolio and is a random variable. These latter two terms depend on the securities the agent chooses to include in the portfolio and hence are subscripted by i . The single index model assumes that \bar{r}_m and \tilde{r}_i are independent, and that the \tilde{r}_i are independent of one another and have a mean of zero.

This model can be put in the form of Model 1 or Model 2 with additional assumptions concerning \tilde{r}_i the nonsystematic risk term. If $\tilde{r}_i = \beta_i \tilde{\delta}_i$ where $\tilde{\delta}_i$ are identically distributed, then the single index model becomes:

$$\text{Model 1A: } \bar{r}_i = \phi + \beta_i(\bar{r}_m - \phi + \tilde{\delta}_i)$$

where the various terms satisfy the restrictions contained in Model 1, including the requirement that the random parameter, $(\bar{r}_m - \phi + \tilde{\delta}_i)$ be identically distributed for each \bar{r}_i .

Similarly, if the \tilde{r}_i terms, when scaled to have the same variance, are identically distributed, then a model like Model 2 results. Let θ_i denote a scale factor so that $\tilde{\phi}_i = \theta_i \tilde{\delta}_i$ where $\tilde{\delta}_i$ are identically distributed. The single index model then takes the form

$$\text{Model 2A: } \tilde{r}_i = \phi + \beta_i (\tilde{r}_m - \phi) + \theta_i \tilde{\delta}_i$$

which satisfies the form and other restrictions of Model 2. Thus, if the nonsystematic risk variables, when scaled appropriately, are identically distributed then the EU efficient set is contained in the MS efficient set.

In the CAPM model all terms other than the nonsystematic risk term are observable or can be obtained from independent sources. Hence the test we conduct focuses on whether or not the calculated δ_{it} are likely to have resulted as random samples from identically distributed $\tilde{\delta}_i$. That is, using

$$\text{Model 1A: } \delta_{it} = [r_{it} - \phi - \beta_i (r_{mt} - \phi)] / \beta_i \quad \text{or}$$

$$\text{Model 2A: } \delta_{it} = [r_{it} - \phi - \beta_i (r_{mt} - \phi)] / \theta_i$$

and data on the terms on the right hand side, one can test whether or not the calculated $\{\delta_{it}\}$ are likely to have resulted from draws from the same population. Since theory suggests that the $\tilde{\delta}_i$ are independent of one another, the Kolmogorov-Smirnov (KS) or Cramer-von-Mises tests are appropriate for this problem.

The KS test is selected because it is easy to compute, and hence easily extendable to cases of many samples and many observations. For portfolios of common stock, a large number of nonoverlapping samples and observations are available. Since the power of these statistical tests increases as the number of samples or the sample size increases, it is advantageous to be able to deal with large numbers of samples

with many observations.

The multisample Kolmogorov-Smirnov (KS) statistic, denoted D , is the maximum difference between any pair of k empirical distribution functions (EDFs) formed from k samples. Formally, D is given by $D = \sup |F_i(x) - F_j(x)|$, where the supremum is taken across x , i , and j , and $F_i(x)$ and $F_j(x)$ are the EDFs from samples i and j . The statistic D is a discrete random variable which takes on a rational value between zero and one.

The probability distribution for the statistic D has been analytically derived and tabulated for the two and three sample cases under the assumption that the samples are independently drawn from the same population. It has also been tabulated for a larger number of samples using Monte Carlo methods. This is possible because the probability distribution for D does not depend on the form of the continuous distribution function describing the population and D is easily computed.³

For the two sample case, the KS test compares favorably with other nonparametric tests and with parametric tests of this same hypothesis. The main nonparametric alternative is the Cramer-von Mises test. The parametric alternatives vary, depending on the form of the distribution function assumed and the specific alternative hypothesis. Conover [1971] indicates that the KS test is similar in power to the Cramer-von Mises test, easier to compute, and is more extensively tabulated, especially for small samples. Compared with parametric tests, such as the Chi-square or Lillifors tests which assume normality, the KS test does quite well, with only a small loss in power under normality. The KS test does much better than these

parametric tests when normality is violated, and hence these latter tests are misapplied.

The distribution free property of the statistic D is also a desirable feature since neither Model 1 or 2 make any statement concerning the form of the distribution functions describing the random variables. Bradley [1968] indicates that the KS test is sensitive to all alternative hypotheses, not just ones concerning differences in location or scale. That is, it is a test which is also sensitive to differences in the shape or form of the distribution function.

IV. Portfolio Data and Hypothesis Test Results

In this section, results from using the multisample KS test to test the hypothesis of identically distributed scaled nonsystematic risk in the rate of return data for portfolios of common stock are reported. All the data used in the study are drawn from the CRSP⁴ tapes of monthly rate of returns on corporate equities traded on the New York or American Stock Exchanges. The rate of return includes dividends and capital gains. These tapes were searched to find all securities for which a thirty year history of monthly data (360 observations) exists over the period January, 1955 through December, 1984. A total of 424 securities were found.

This sample was split into two randomly chosen subsamples of 212 securities each.⁵ One of these subsamples was reserved for future analysis, and has not been used to date. A detailed examination of the second subsample revealed 20 securities for which one or more monthly rates of return are missing. These securities were dropped

from the sample and not replaced. Thus, the full working sample which is used to build portfolios has 360 monthly observations on 192 separate securities.

This full working sample on the 192 securities was split once again into two randomly chosen subsamples of the same 180 monthly observations on each security. Subsample A is used to independently obtain values for the parameters β_1 and θ_1 . The only use made of subsample A is to construct values for these two constants. Subsample B is used to test for identically distributed nonsystematic risk when the nonsystematic risk variable is scaled by one of the constants determined from subsample A. A more detailed description of how this is accomplished follows.

Using the 180 monthly observations in subsample A, the rates of return for each of the 192 securities are regressed against the rate of return for the market portfolio for those same months. The market rate of return is represented by the CRSP market index which uses value weighting and includes dividends. The regression slope coefficient for each security is used as its β_1 . The standard deviation of the residuals in each regression is used as the θ_1 for that security.

To construct portfolios, the 192 securities were numbered from 1 to 192, and a securities sampling vector was constructed containing those numbers ordered by the corresponding security's β_1 value. When tests involve n portfolios of k securities in each, the i th portfolio contains the securities indicated by elements $(i-1)(190/n) + j$ for $j = 1$ to k of this securities sampling vector. In our analysis n always is either 10, 19 or 38.

Given the sample of 192 securities, this procedure allows one to create 10 (19, 38) portfolios containing up to 19 (10, 5) securities with no two portfolios of the same size containing common securities. This no-overlap condition allows independent nonsystematic risk terms for the various portfolios of a fixed size. Under this procedure, as the number of securities included in the portfolios increases the previous securities are retained and new securities are added. In each case investigated, the rate of return on a multiple security portfolio is an equal weight ($1/k$) average of the rates of return of the component securities. The β_i and θ_i for the portfolios are determined in the same manner as for securities. This portfolio building procedure gives portfolios with widely differing β_i values, both efficient and inefficient ones.⁶

Given values for β_i and θ_i , a constant risk free rate (ϕ) of 6% per year (rate of return of .005 per month), and the observed market rate of return r_{mt} , the observations on \tilde{r}_i in subsample B are used to calculate sets of values for the unobserved scaled nonsystematic risk under Model 1A or 2A. This is done for each portfolio using the transformations mentioned earlier and relisted below for convenience.

$$\text{Model 1A: } \delta_{it} = [r_{it} - \phi - \beta_i(r_{mt} - \phi)]/\beta_i$$

$$\text{Model 2A: } \delta_{it} = [r_{it} - \phi - \beta_i(r_{mt} - \phi)]/\theta_i$$

The question is: are these sets of values likely to have resulted as random samples from the same population.

These sets of values initially contain 180 points. To investigate subsamples of smaller size, an observations sampling vector consisting of 180 random numbers between 1 and 180, drawn without replacement, was constructed. When tests are conducted using samples of size k ,

the observations whose numbers match the first k elements of this observations sampling vector are used. Thus, as the length of the sample used in the statistical test is increased, the observations from the shorter samples are retained and additional observations are merged with them.

When k portfolios of the same number of securities are obtained using the above procedures, the method of portfolio construction allows the nonsystematic risk terms to be independent of one another. Thus, the KS test with k samples and n observations is appropriate. These tests were conducted for k = 10, 19 and 38 samples (portfolios) and using n = 25, 50, 100 and 180 observations. In addition, the number of securities included in the portfolios ranged from 1 to 19, 10 or 5, depending on whether k = 10, 19 or 38.

A few typical results are listed in Table 1. The entire set of test results are reported in Appendix A. Each element of Table 1 or Appendix A lists the results from the test of a particular hypothesis. This hypothesis is that the set of values representing scaled nonsystematic risk obtained using either Model 1A or 2A with the indicated parameter values, are random samples from the same population.

Each line in the table or the Appendix contains four test results. They are for the indicated model, number of portfolios, and number of securities in the portfolio, and for samples of 25, 50, 100 and 180 observations. The numbers reported are the observed value for D as a fraction (12/25), and below it a pair of numbers (.93, .16) giving the probability of observing a D value less than or equal to, or greater than or equal to the observed value, respectively. Neither of these

probabilities is likely to be small if the sets of data are independent samples from the same population.⁷

Table 1 gives the test results for both models for the 10 portfolio case and for portfolios containing 1, 5, 10, and 15 securities. These results are typical of those listed in Appendix A. The entry for portfolios with 15 securities and 180 observations indicates that when 10 portfolios are selected with 15 different securities in each and 180 monthly observations are obtained, the scaled nonsystematic risk terms for these 10 portfolios are not likely to be identically distributed when scaled by $1/\beta_1$ and thus model 1A is rejected. This conclusion is indicated by the KS statistic's value 39/180. This large a value only occurs 2% of the time if the samples are from the same population.

On the other hand, if the nonsystematic risk is scaled by $1/\theta$, the identical distribution hypothesis is not rejected (Model 2A). In this case the observed value for D is 28/180 which is neither too large nor too small as to be unlikely under the identical distribution hypothesis. The probability of observing this value or less is .47, while the probability of this value or greater is .61. The remaining entries in these tables of results are interpreted in this same fashion.

Two observations can be made concerning the test results. First, when the number of observations is either 25 or 50, neither model can be rejected. One finds very few instances where either probability is smaller than .10 or even .20 and certainly no more than one would expect assuming these rejection levels. Second, with larger sample sizes, 100 or 180 observations, Model 1A is quite clearly rejected⁸

and Model 2A is not. Model 1A is rejected at even the .05 level in the vast majority of cases, and rejected in every instance at the .20 level. Model 2A on the other hand is not rejected in most instances at the .05 level and not more often than one would expect at higher significance levels.

Two minor points of information concerning this work are noteworthy. First, the reason why portfolios containing different numbers of securities were selected for study was the prior hypothesis that well diversified portfolios were more likely to satisfy the conditions of Model 1 or 2 than single securities. While rejection is more frequent in portfolios with few securities, Model 1A is rejected even in portfolios with as many as 19 securities, and Model 2A is not rejected even in cases where the number of securities is quite small. Thus, the evidence does not support this prior hypothesis in any convincing way.

Second, the monthly data can be aggregated into quarterly or semi-annual observations and the study repeated. This was carried out in part to see if the holding period had any significant effects on the outcome of these tests. None were found, and the reduced number of observations makes the tests less powerful, so this direction of extension was abandoned.

Conclusions

In this work, estimation error is considered in the tests which are conducted. The tests ask whether or not the random variables whose distribution functions are being estimated are sufficiently similar to one another when estimation error is taken into consideration. Other studies comparing the efficient sets under

EU or MS ranking procedures have used data to estimate the distribution functions for rate of return on portfolios and then treated those estimates as if they are exact. Levy and Markowitz [1979]. Kroll, Levy and Markowitz [1984] and Pulley [1981] are recent examples. Since these researchers are primarily addressing the question of potential differences between the EU and MS ranking techniques they treat their estimates as a means of obtaining representative distribution functions. This use of estimates as if they are exact is appropriate.

The fact that estimation error is taken into account in our procedure means that the finding that Model 2A is not rejected should be interpreted as implying that the MS and EU efficient sets do not differ from one another in a statistically significant way. This also gives added meaning to the often made criticism of EU ranking procedures that they "require more data". It is certainly true that with few enough observations there is no statistically significant difference between EU and MS ranking procedures in that the random alternatives are being estimated so imprecisely so as to not allow rejection of the LS or similar conditions.

This work shows how one can analyze the data to determine if EU ranking methods can possibly give results that are statistically significantly different from MS procedures. Since mutual funds are portfolios, this work indicates that comparing stochastic dominance with MS rankings of these funds on the basis of 10 observations cannot result in statistically significant findings, but can only illustrate the technique involved (Joy and Porter, [1974]; Meyer [1977]).

Rejection of Model 1A or 2A, with the specified parameter values does not imply that Model 1 or Model 2 is rejected. That is, it could be that the rate of return on portfolios of common stock follow a model of the appropriate form, but the test carried out chose the incorrect random parameter \tilde{y}_1 , or incorrect values for a_1 and b_1 . On the other hand, if the specific version of the model is not rejected, then there exist some parameters such that a model of the appropriate form represents the \tilde{z}_1 adequately.

In the specific implementation of Model 1 and 2 used here, the values for the parameters, and the market index selected to represent the common source of randomness, are picked using the suggestions of the the very simplest CAPM. More sophisticated methods of obtaining values for these parameters are available, other market indices could be selected, and some evidence even suggests that certain of the parameters which are assumed to be constant are not fixed over time. Any of these changes can lead to better models explaining rate of return on portfolios, or better prediction of future values for $r_{i,t}$, but cannot lead to rejection of Model 2. That is, this work has indicated that there exist constants a_1 and b_1 and random variables \tilde{x}_1 and \tilde{y}_1 such that the form and restrictions listed as Model 2 could reasonably be assumed to lead to the data observed concerning rate of return on stock portfolios. Only new data sets or better tests can alter this conclusion.

Footnotes

1. The model of the competitive firm facing a random output price of Sandmo [1971] and the portfolio model of Tobin [1958] are examples displaying this structure.

2. With independence of \tilde{x}_1 and \tilde{y}_1 , the following two equations relate a_1 and b_1 to μ_1 and σ_1 .

$$\mu_1 = (a_1^2 \mu_y^2 + b_1^2 \mu_x^2)^{1/2} = T^{1/2} \quad (1)$$

$$\sigma_1 = a_1 \sigma_y + b_1 \sigma_x + c \quad (2)$$

These equations can be inverted and used to determine μ and σ from a and b if the determinant of H :

$$H = \begin{vmatrix} T^{-1/2}(a_1 \mu_y^2) & T^{-1/2}(b_1 \mu_x^2) \\ \sigma_y & \sigma_x \end{vmatrix}$$

is not zero. Assuming this is the case, define $V(\mu, \sigma)$ to be the ranking function over (μ, σ) space which represents the expected utility preferences over alternatives \tilde{z}_1 , when the \tilde{z}_1 result from a model with this general structure. It is now a straightforward calculation to develop the properties of $V(\mu, \sigma)$ and relate them to properties of $u(z)$. One can show that the sign of V_σ depends on risk aversion in the expected utility model. Risk aversion in the expected utility model implies that the agent always chooses that alternative with the lowest standard deviation from among those with a given mean value.

3. This work has been carried out to a limited extent by Gardner, Pinder and Wood [1980]. Since their work does not examine enough samples or samples of sufficient size for our purposes, Appendix B describes the Monte Carlo procedures used to obtain the small sample distribution for D for any number of samples of any size.

4. CRSP monthly stock return tape is maintained by the Center for Research in Security Prices, University of Chicago.

5. The "IPERM" numbers of these firms are available from the authors upon request.

6. We also conducted tests where the portfolios contained randomly selected securities without regard to β_1 . For portfolios containing several securities or more, the rate of returns distributions become similar to one another not only in form, but also in level of mean and variance. Hence being able to use a mean-variance ranking procedure is not very valuable. The procedure reported in the text ensures that the portfolios are representative of the broad spectrum of portfolios which can be formed. Inefficient as well as efficient ones are included, and the portfolio's β_1 values range from about .5 to 2.0.

7. The sum of these two probabilities minus one is the probability of obtaining the observed value of D . At one point in the analysis we accidentally constructed portfolios which contained securities in common. In some instances an overlap of about 50% occurred. This

caused the observed KS values to be too small. It appears that positive correlation leads to lower values for D than would occur with independence. It is for this reason that we report the probability that the D value is less than or equal to the one observed.

8. We also tried a version of Model 1 in which the risk free rate is replaced by the intercept of the regression used to obtain the b_i values. This did not change the conclusions concerning Model 1.

References

- Birnbaum Z.W. and Hall, R.A., "Small Sample Distributions For Multi-Sample Statistics of the Smirnov Type," Annals of Mathematical Statistics, 1960, 31, 710-20.
- Bradley, J.V., Distribution-Free Statistical Tests, Prentice-Hall, Englewood Cliffs, N.J., 1968.
- Conover, W., Practical Nonparametric Statistics, John Wiley and Sons Inc., New York, 1971.
- Feder, G., "The Impact of Uncertainty in a Class of Objective Functions," J. Econ. Theory, 1977, 16, 504-12.
- Gardner, R., Pinder, J., and Wood, R., "Monte Carlo Estimation of Percentiles for the Multi-Smirnov Test," J. Statist. Comput. Simul., 1980, 10, 243-49.
- Joy, O. M., and Porter, R. B., "Stochastic Dominance and Mutual Fund Performance", J. of Financial and Quantitative Analysis, 1974, 9, 25-31.
- Kroll, Y., Levy, H. and Markowitz, H., "Mean-Variance Versus Direct Utility Maximization," Journal of Finance, 1984, 39, 47-62.
- Levy, H. and Markowitz, H., "Approximating Expected Utility by a Function of Mean and Variance," Amer. Econ. Rev., 1979, 69, 308-17.
- Manoukian, E.B., Mathematical Nonparametric Statistics, Gordon and Breach Science Publishers, New York, 1986.
- Meyer, J., "Further Applications of Stochastic Dominance to Mutual Fund Performance", J. of Financial and Quantitative Analysis, 1977, 12, 235-242.
- Meyer, J., "Two Moment Decision Models and Expected Utility Maximization," Amer. Econ. Rev., 1987, 77, 421-30.
- Pulley, L.M., "A General Mean-Variance Approximation to Expected Utility for Short Holding Periods," Journal of Financial and Quantitative Analysis, 1981, 16, 361-73.
- Rothschild, M. and Stiglitz, J., "Increasing Risk: 1. A Definition," J. Econ. Theory, 1970, 2, 225-43.
- Sandmo, A., "On The Theory of the Competitive Firm Under Price Uncertainty," Amer. Econ. Rev., 1971, 37, 537-42.
- Sinn, Hans-Werner, Economic Decisions Under Uncertainty, North-Holland Pub. Co., New York, 1983.

Taylor, M.S. and Becker, M.F., "Quantiles of Multi-Sample Smirnov Type Statistics," J. Statist. Comput. Simul., 1982, 16, 25-34.

Tobin, J., "Liquidity Preference as Behavior Towards Risk," Review of Economic Studies, 1958, 25, 65-86.

Table 1

10 PORTFOLIOS: Model 1A

Number of Securities	Number of Observations			
	25	50	100	180
1	12/25 (.93, .16)	17/50 (.92, .16)	29/100 (.99, .02)	34/180 (.93, .10)
5	11/25 (.84, .34)	16/50 (.84, .27)	24/100 (.89, .17)	40/180 (.99, .01)
10	12/25 (.93, .16)	16/50 (.84, .27)	34/100 (.99, .00)	51/180 (.99, .00)
15	10/25 (.66, .57)	16/50 (.84, .27)	25/100 (.93, .11)	39/180 (.99, .02)

10 PORTFOLIOS: Model 2A

Number of Securities	Number of Observations			
	25	50	100	180
1	12/25 (.93, .16)	16/50 (.84, .27)	28/100 (.98, .02)	31/180 (.81, .24)
5	9/25 (.43, .82)	15/50 (.73, .42)	21/100 (.67, .44)	25/180 (.34, .74)
10	10/25 (.66, .57)	13/50 (.40, .73)	24/100 (.89, .17)	36/180 (.97, .05)
15	10/25 (.66, .57)	14/50 (.58, .60)	20/100 (.56, .57)	28/180 (.61, .47)

Appendix A

10 PORTFOLIOS: Model 1A

Number of Securities	Number of Observations			
	25	50	100	180
1	12/25 (.93, .16)	17/50 (.92, .16)	29/100 (.99, .02)	34/180 (.93, .10)
2	9/25 (.43, .82)	15/50 (.73, .42)	24/100 (.89, .17)	32/180 (.86, .19)
3	10/25 (.66, .57)	15/50 (.73, .42)	25/100 (.93, .11)	37/180 (.98, .03)
4	11/25 (.84, .34)	14/50 (.58, .60)	27/100 (.98, .04)	39/180 (.99, .02)
5	11/25 (.84, .34)	16/50 (.84, .27)	24/100 (.89, .17)	40/180 (.99, .01)
6	11/25 (.84, .34)	16/50 (.84, .27)	25/100 (.93, .11)	40/180 (.99, .01)
7	11/25 (.84, .34)	15/50 (.73, .42)	29/100 (.99, .02)	37/180 (.98, .03)
8	12/25 (.93, .16)	15/50 (.73, .42)	28/100 (.98, .02)	41/180 (.99, .01)
9	12/25 (.93, .16)	15/50 (.73, .42)	31/100 (.99, .01)	47/180 (.99, .00)
10	12/25 (.93, .16)	16/50 (.84, .27)	34/100 (.99, .00)	51/180 (.99, .00)
11	12/25 (.93, .16)	16/50 (.84, .27)	33/100 (.99, .00)	49/180 (.99, .00)
12	12/25 (.93, .16)	14/50 (.58, .60)	29/100 (.99, .02)	47/180 (.99, .00)
13	11/25 (.84, .34)	15/50 (.73, .42)	27/100 (.98, .04)	46/180 (.99, .00)
14	11/25 (.84, .34)	16/50 (.84, .27)	26/100 (.96, .07)	42/180 (.99, .00)
15	10/25 (.66, .57)	16/50 (.84, .27)	25/100 (.93, .11)	39/180 (.99, .02)
16	10/25 (.66, .57)	16/50 (.84, .27)	27/100 (.98, .04)	42/180 (.99, .00)
17	11/25 (.84, .34)	15/50 (.73, .42)	28/100 (.98, .02)	42/180 (.99, .00)
18	12/25 (.93, .16)	15/50 (.73, .42)	27/39 (.98, .04)	39/180 (.99, .02)
19	10/25 (.66, .57)	15/50 (.73, .42)	28/100 (.98, .02)	40/180 (.99, .01)

10 PORTFOLIOS: Model 2A

Number of Securities	Number of Observations			
	25	50	100	180
1	12/25 (.93, .16)	16/50 (.84, .27)	28/100 (.98, .02)	31/180 (.81, .24)
2	10/25 (.66, .57)	13/50 (.40, .73)	25/100 (.93, .11)	31/180 (.81, .24)
3	10/25 (.66, .57)	13/50 (.40, .73)	25/100 (.93, .11)	25/180 (.34, .74)
4	11/25 (.84, .34)	12/50 (.23, .90)	23/100 (.83, .24)	23/180 (.18, .88)
5	9/25 (.43, .82)	15/50 (.73, .42)	21/100 (.67, .44)	25/180 (.34, .74)
6	11/25 (.84, .34)	17/50 (.92, .16)	22/100 (.76, .33)	28/180 (.61, .47)
7	9/25 (.43, .82)	15/50 (.73, .42)	29/100 (.99, .02)	34/180 (.93, .10)
8	11/25 (.84, .34)	14/50 (.58, .60)	27/100 (.98, .04)	30/180 (.76, .30)
9	11/25 (.84, .34)	12/50 (.23, .90)	26/100 (.96, .07)	35/180 (.95, .07)
10	10/25 (.66, .57)	13/50 (.40, .73)	24/100 (.89, .17)	36/180 (.97, .05)
11	11/25 (.84, .34)	12/50 (.23, .90)	24/100 (.89, .17)	31/180 (.81, .24)
12	10/25 (.66, .57)	12/50 (.23, .90)	23/100 (.83, .24)	32/180 (.86, .18)
13	8/25 (.18, .95)	11/50 (10., .97)	21/100 (.67, .44)	28/180 (.61, .47)
14	9/25 (.43, .82)	12/50 (.23, .90)	17/100 (.19, .90)	23/180 (.18, .88)
15	10/25 (.66, .57)	14/50 (.58, .60)	20/100 (.56, .57)	28/180 (.61, .47)
16	8/25 (.18, .95)	16/50 (.84, .27)	18/100 (.31, .81)	28/180 (.61, .47)
17	9/25 (.43, .82)	14/50 (.58, .60)	20/100 (.56, .57)	30/180 (.76, .30)
18	9/25 (.43, .82)	15/50 (.73, .42)	19/100 (.43, .69)	29/180 (.70, .39)
19	9/25 (.43, .82)	15/50 (.73, .42)	20/100 (.56, .57)	29/180 (.70, .39)

19 PORTFOLIOS: Model 1A

Number of Securities	Number of Observations			
	25	50	100	180
1	13/25 (.93, .18)	19/50 (.93, .13)	36/100 (.99, .00)	63/180 (.99, .00)
2	13/25 (.93, .18)	18/50 (.87, .24)	33/100 (.99, .01)	52/180 (.99, .00)
3	13/25 (.93, .18)	18/50 (.87, .24)	32/100 (.99, .01)	50/180 (.99, .00)
4	11/25 (.61, .65)	14/50 (.23, .90)	30/100 (.98, .03)	48/180 (.99, .00)
5	13/25 (.93, .18)	17/50 (.76, .39)	29/100 (.97, .05)	47/180 (.99, .00)
6	12/25 (.83, .39)	18/50 (.87, .24)	31/100 (.99, .02)	50/180 (.99, .00)
7	11/25 (.61, .65)	17/50 (.76, .39)	33/100 (.99, .01)	49/180 (.99, .00)
8	12/25- (.83, .39)	14/50 (.23, .90)	30/100 (.98, .03)	45/180 (.99, .00)
9	13/25 (.93, .18)	17/50 (.76, .39)	29/100 (.97, .05)	43/180 (.99, .01)
10	12/25 (.83, .39)	16/50 (.61, .58)	29/100 (.97, .05)	40/180 (.98, .03)

38 PORTFOLIOS: Model 1A

Number of Securities	Number of Observations		
	25	50	100
1	14/25 (.93, .20)	20/50 (.90, .20)	39/100 (.99, .00)
2	14/25 (.93, .20)	19/50 (.80, .35)	34/100 (.99, .01)
3	13/25 (.80, .46)	18/50 (.65, .55)	35/100 (.99, .01)
4	12/25 (.54, .74)	15/50 (.10, .97)	34/100 (.99, .01)
5	13/25 (.80, .46)	17/50 (.45, .75)	30/100 (.95, .09)

19 PORTFOLIOS: Model 2A

Number of Securities	Number of Observations			
	25	50	100	180
1	12/25 (.83, .39)	18/50 (.87, .24)	26/100 (.87, .21)	33/180 (.74, .34)
2	14/25 (.98, .07)	17/50 (.76, .39)	26/100 (.87, .21)	34/180 (.80, .26)
3	13/25 (.93, .18)	19/50 (.93, .13)	30/100 (.98, .03)	36/180 (.90, .15)
4	11/25 (.61, .65)	14/50 (.23, .90)	21/100 (.32, .79)	29/180 (.37, .72)
5	11/25 (.61, .65)	15/50 (.42, .73)	22/100 (.45, .68)	31/180 (.58, .52)
6	12/25 (.83, .39)	16/50 (.61, .58)	26/100 (.87, .21)	34/180 (.80, .26)
7	10/25 (.35, .88)	15/50 (.42, .73)	24/100 (.70, .42)	31/180 (.58, .52)
8	13/25 (.93, .18)	14/50 (.23, .90)	22/100 (.45, .68)	29/180 (.37, .72)
9	11/25 (.61, .65)	16/50 (.61, .58)	22/100 (.45, .68)	32/180 (.66, .42)
10	9/25 (.12, .98)	14/50 (.23, .90)	23/100 (.58, .55)	27/180 (.19, .88)

38 PORTFOLIOS: Model 2A

Number of Securities	Number of Observations		
	25	50	100
1	13/25 (.80, .46)	19/50 (.80, .35)	27/100 (.76, .35)
2	14/25 (.93, .20)	18/50 (.65, .55)	31/100 (.97, .05)
3	13/25 (.80, .46)	19/50 (.80, .35)	30/100 (.95, .09)
4	11/25 (.26, .94)	14/50 (.03, .99)	24/100 (.36, .77)
5	11/25 (.11, .94)	15/50 (.10, .97)	24/100 (.36, .77)

Appendix B

Monte Carlo Methods and Small Sample Distributions for the Kolmogorov-Smirnov Statistic

Several studies compute the exact values of the cumulative distribution function for the KS statistic D , but do so only for small numbers of samples. Birnbaum and Hall [1960] focus on the two and three sample cases, and Taylor and Becker [1982] extend their work to the four sample case and also allow unequal sample sizes. Unfortunately, the computations required for the calculation of these exact values get prohibitively expensive as either the number of samples or the number of observations becomes large.

An alternative approach is to estimate the cumulative distribution function for D using Monte Carlo methods. This approach was utilized by Gardener, Pinder and Wood [1980] who give percentiles for D for sample sizes up to 100 and for up to 10 samples. They conclude that accurate approximations to the known exact distribution functions for the two and three samples cases can be obtained using 5000 replications in the Monte Carlo experiment.

This Monte Carlo approach is adaptable to currently available personal computer technology. A small program in the GAUSS Mathematical and Statistical Programming Language¹ was written to compute efficiently the value for D for arbitrary numbers of samples and sample sizes, and to use these values from a large number of replications to estimate the CDF for D . The source code for this program and a more detailed description of the algorithm are in Appendix C. The program requires equal sample size across the various samples for reasons discussed below. This restriction is of no consequence for the types of economic and financial data dealt with here.

The accuracy of this particular Monte Carlo algorithm was examined by estimating the distribution function for D for the three sample case and comparing the results with those computed exactly by Birnbaum and Hall. Following Gardner, Pinder and Wood, 5000 replications were used. The results confirm that the procedure is highly accurate. There are 234 entries in the Birnbaum and Hall Table 1 ranging from 0 to 1, but with almost all being larger than .5. The mean error of our estimates of these values is .0004, with a standard deviation of .0038. The estimation errors are distributed quite uniformly across the various entries in the Birnbaum and Hall table. The maximum error observed was .0134. Thus, we are confident the algorithm given in Table C1 with 5000 replications estimates the distribution function for D with error whose magnitude is of no consequence given the critical values (.05 - .2) employed in this study.

The algorithm allows the user to specify the number of samples (k), the number of observations (n), and the number of replications (r) to carry out. The time required to compute the estimated cumulative distribution for D is not excessive as long as the equal sample size restriction is imposed. Some examples of the computational time required to carry out 50 replications for various sample sizes and number of observations are given in Table B1². These times are such that .0025 times the product of the number of samples, observations and replications (.0025 knr) gives a good approximation

to the total time in seconds required to obtain the distribution function for D for the specified case. For the 5000 replication case this is approximately $(nk)/300$ hours.

The reason that equal sample size improves the speed of the calculation can be seen in Figure I. To compute the value for D, one needs the maximum distance between the upper and lower envelopes of the EDFs formed from the k samples. Finding these envelopes and the value for D involves searching across all of the nk observations since the envelopes can change value at any of them. Since the EDFs are nondecreasing functions however, the upper and lower envelopes are also the left and right envelopes. Under the equal sample sized restriction these are easier to find since they can change values only at n different points.³ Thus, the problem is reduced from one of dimension kn to one of dimension $2n$. Sample sizes (k) equal to values up to 180 are dealt with here so this proves to be a significant shortcut. More details concerning this algorithm are included in Appendix C.

Certain observations concerning the distribution function for D are worth noting. First, the distribution for D becomes more concentrated about its mean as either the sample size or the number of samples increases, but this concentration is much more sensitive to sample size than number of samples. Test statistics for which the probability mass is more concentrated about the mean value are more useful in hypothesis testing. Thus, the analysis here pays more attention to the effects of sample size than number of samples.

Second, for the two sample case, the limiting distribution for D, as the number of observations increases, has been derived and involves the scale factor $n^{1/2}$; that is, $n^{1/2}D$ has a known limiting distribution. This limiting distribution is a good approximation for sample sizes of 75 - 100 or larger. (Manoukian [1986]) Extensions of this finding to cases involving more than two samples have not been found. Some evidence, however, is given in Figures II and III. In Figure II, smoothed histograms for D for the 2, 10, 19 and 38 sample cases are given in four different panels. Various numbers of observations per sample are included. These are scaled so each contains unit area. It is clear from this figure that the cases involving a larger number of samples behave qualitatively like the two sample case. Indeed, when the D values are transformed by the scale factor $n^{1/2}$ to obtain the four panels in Figure III, it is clear that a limiting distribution exists for those cases as well. Again samples sizes of 75 - 100 or larger appear to be sufficient to use the limiting distribution.

One might also ask if there is a limiting distribution as k, the number of samples, increases. From Figure IV one can observe that a scale factor in k must be used and that it must be nonlinear in k. So far, we have been unable to approximate such a factor as a simple nonlinear function of k.

Footnotes

1. GAUSS, copyright by Aptech Systems, Inc. Box 6487, Kent, WA 98064.
2. These timing statistics were produced on a Zenith 158 PC with an 8mhz clock speed and an 8087 numeric coprocessing chip. Several tests on a Zenith 241 with an 80287 coprocessor indicated a reduction in the

required time by a factor of about 2.5.

3. The right and left envelopes continue to be easier to compute even when sample sizes differ. An algorithm which handles unequal sample size is available from R. Rasche.

Table B1

Examples of Time Requirements for D Statistic
 Monte Carlo Computations
 (Time in Seconds)
 (50 Replications)

n	k	3	10	25	50	100
25		24.17	43.61	85.91	156.04	295.77
50		40.70	77.95	158.30	292.76	606.21
100		78.33	156.09	324.11	605.11	na
150		120.51	242.39	950.48	na	na

na = not available. The implementation of the program given in Table 1 requires that $n*k \leq 8190$.

FIGURE I

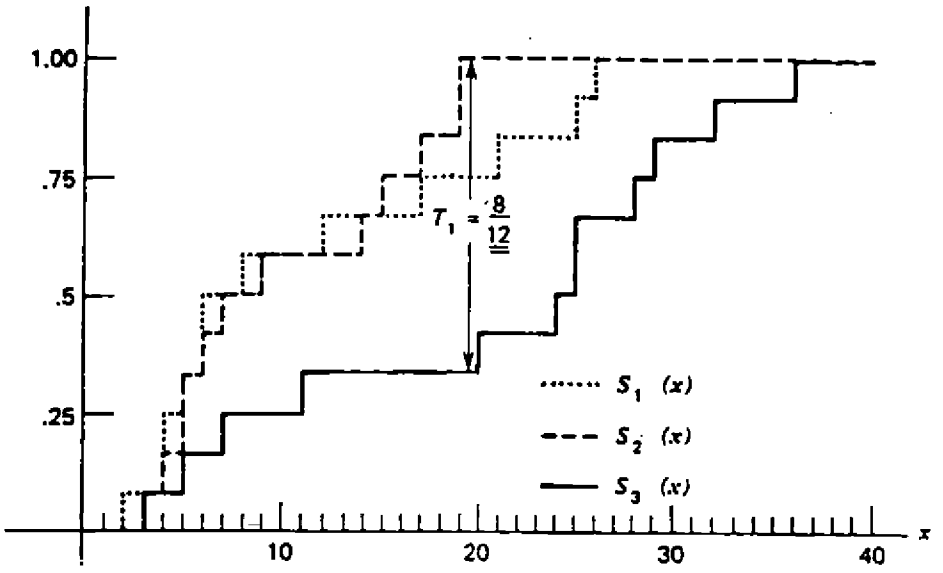


FIGURE II

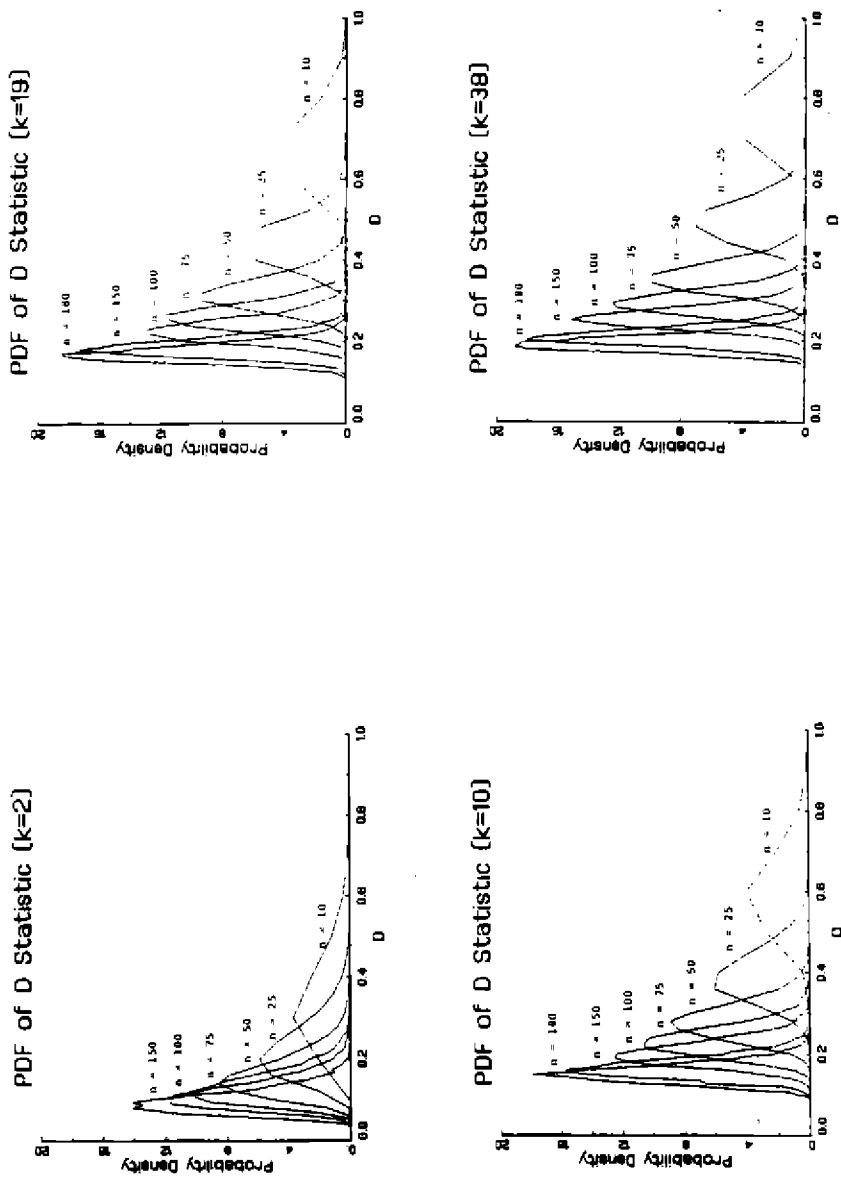


FIGURE III

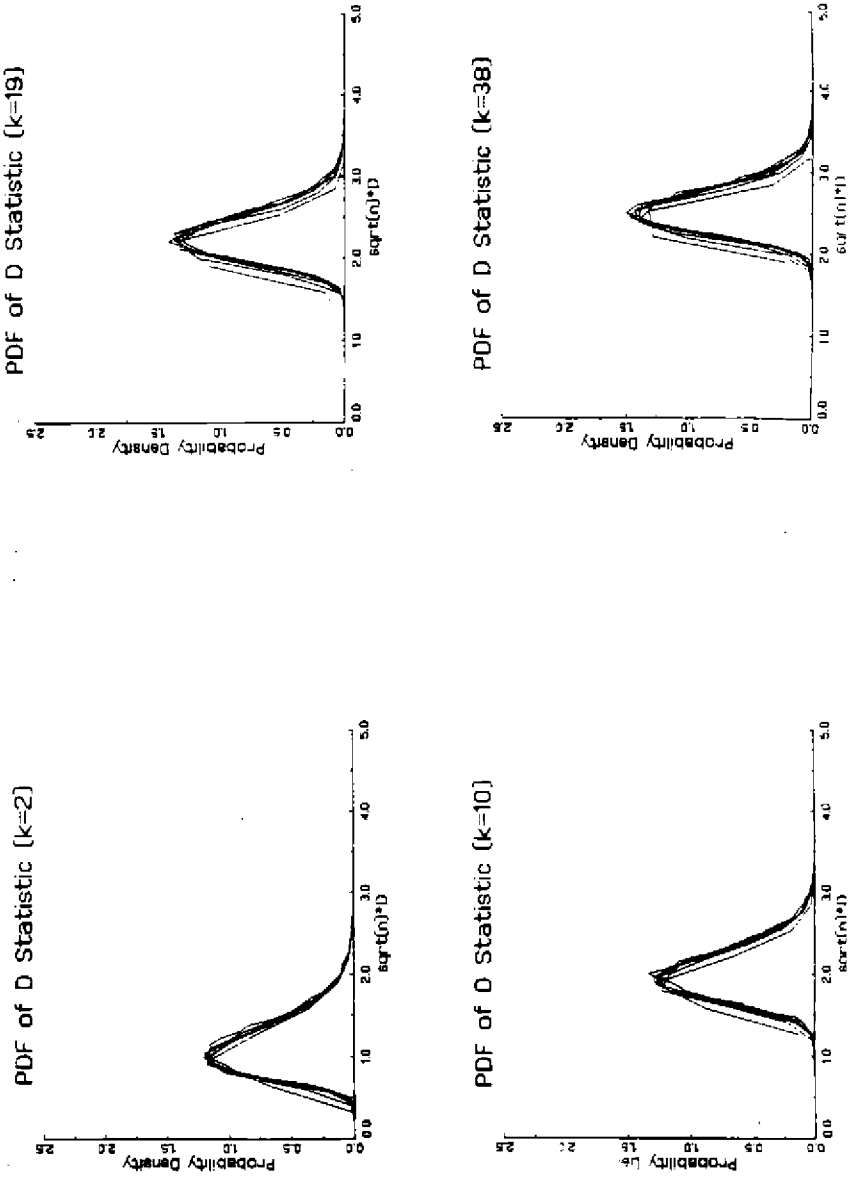
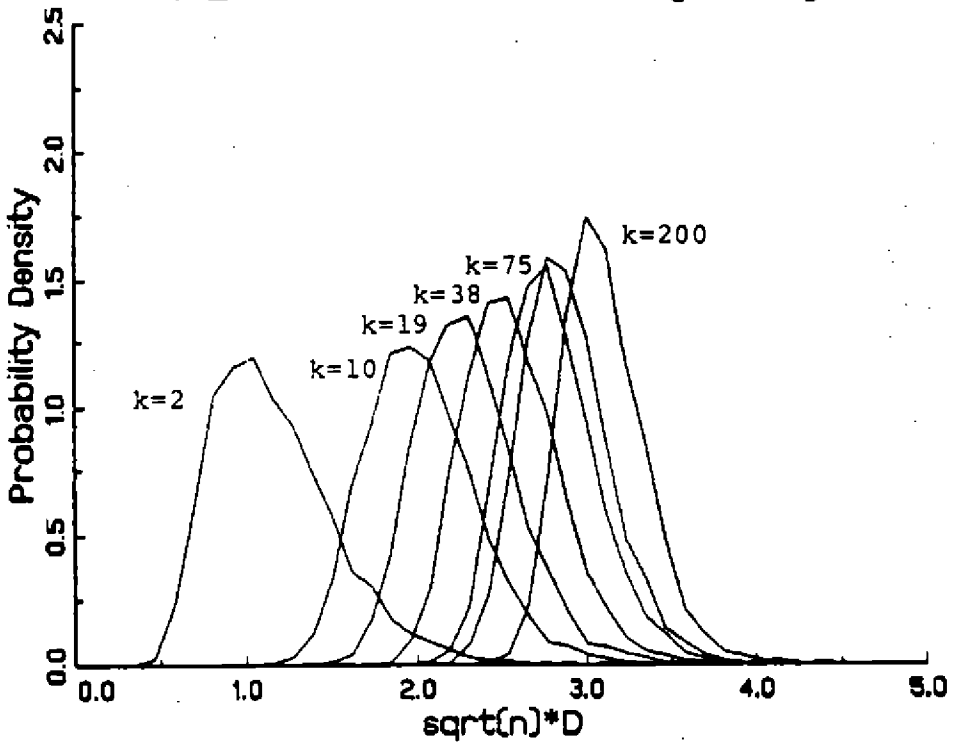


FIGURE IV

PDF of D Statistic (n=75)



Appendix C

The source code in the GAUSS programming language is listed in Table A1. The portion of the algorithm which determines the value for D is illustrated using a small data set involving three samples taken from Conover [p.318, 1971]. The data and the steps of the algorithm are given in Table C2. Each of the steps described below are indicated by a corresponding number in that table. GAUSS is a matrix processing language and the algorithm takes advantage of this fact.

- 1) construct a matrix of k columns of the n observations on the random variables: (x_{ij})
- 2) sort each column of this matrix into ascending order.
- 3) construct a $n \times 2$ matrix, the first column containing the maximum values of each row of the sorted matrix in 2) (\max over j of x_{ij}), and the second column containing elements $-1/n$. The information in the first column pertains to the right envelope of the k empirical distribution functions.
- 4) construct a $n \times 2$ matrix, the first column containing the minimum values of each row of the sorted matrix in 2) (\min over j of x_{ij}), and the second column containing elements $1/n$. The information in the first column pertains to the left envelope of the k empirical distribution function.
- 5) construct a $2n \times 2$ matrix by "stacking" the matrix in 3) above the matrix in 4).
- 6) sort the rows of the matrix in 5) into ascending order by their first element.
- 7) compute a vector of the cumulative sums of the elements in the second column of the matrix in 6). These values are the difference between the upper and lower envelopes of the k empirical distribution functions at the 2n points where they change values.
- 8) determine the maximum value of the elements in the vector in 7). This value is the value for D for the k samples.

Table C1

GAUSS Program to Estimate pdf and cdf of the D Statistic

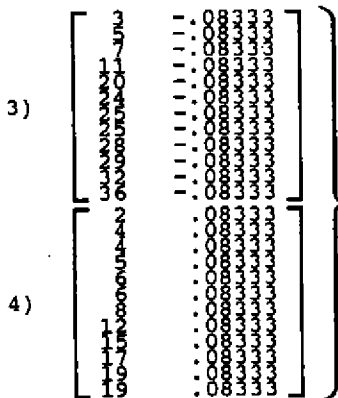
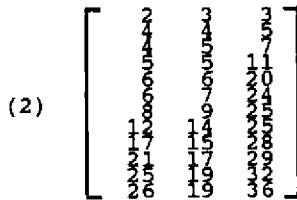
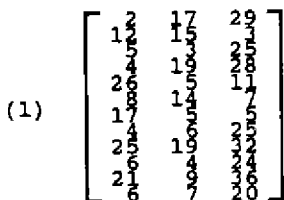
```

print "read num samples, number obs, number iterations";
dims=con(3,1);
lprint "num samples/num observations/num iterations";
lprint /m3 dims;
z=zeros(dims[2,1],dims[1,1]);
down=ones(dims[2,1],2);
up=-ones(dims[2,1],2);
res=zeros(dims[2,1],dims[4,1]+1);
t1=hsec;
j=1;
do until j > dims[3,1];
  cls;
  t2 = (hsec - t1)/100;
  print "iteration number ";;j;;" elapsed time ";;t2;;" seconds";;
  i=1;
  do until i > dims[1,1];
    z[.,i]= sortc(rndu(dims[2,1],1),1);
    i=i+1;
  endo;
  endif;
  down[.,1]=maxc(z');
  up[.,1]=minc(z');
  x=sortc(down|up,1);
  y=recserar(x,x[1,.,],ones(1,2));
  jj = abs(minc(y[.,2]));
  res[jj,1] = res[jj,1] +1;
  j=j+1;
endo;
res = res ./ dims[3,1];
cx=recserar(res[.,1],res[1,1],ones(1,1));
cls;
lprint /m3 "probabilities: ";
lprint /m3 res[.,1];
lprint /m3 "cumulative probabilities";
lprint /m3 cx;
t1 = (hsec-t1)/100;
cls; print "total time: ";;t1;
end;

```

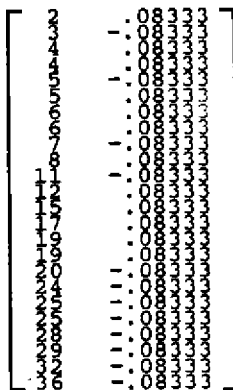
Table C2

Example of Algorithm for Computing D-Statistic with Equal Observations per Sample

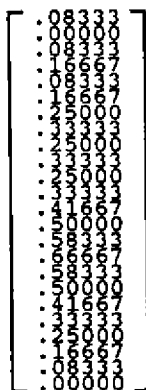


(5)

(6)



(7)



\hat{D}