Dominance Testing in Economics and Finance

Kuan Xu

A Thesis

in

The Department

of

Economics

Presented in Partial Fulfilment of the Requirements

for the Degree of Doctor of Philosophy at

Concordia University

Montreal, Quebec, Canada

June, 1994

National Library of Canada
Acquisitions and Bibliographic Services Branch
395 Wellington Street
Ottawa, Ontario
K1A 0N4

Bibliothèque nationale du Canada
Direction des acquisitions et des services bibliographiques
395, rue Wellington
Ottawa (Ontario)
K1A 0N4

Your file Votre référence

Our file Notre référence

ISBN 0-315-97614-4

Canadä

Abstract

Dominance Testing in Economics and Finance

Kuan Xu, Ph.D.

Concordia University, 1994

This thesis proposes new tests for stochastic dominance that have applications in economics and finance. The existing algorithms and test procedures for stochastic dominance are noticeably restrictive in the sense that (i) they are not firmly based on sampling theory; or (ii) they are restricted to a special class of known parametric distribution functions; or (iii) they do not specify dominance relations properly under the null hypothesis; or (iv) they need restrictive assumptions concerning the data generating processes and hence cannot accommodate complex data structures. This thesis proposes new distribution-free statistical tests which specify the null hypothesis properly, have desirable asymptotic properties, and lead to straightforward methods of inference. The finite sample properties of the proposed tests are evaluated using Monte Carlo simulations. An investigation of the dominance characteristics of real term premium for U.S. Treasury bills is employed to illustrate the application of these tests.

Acknowledgements

There are so many individuals to whom I feel grateful. They have made my stay at Concordia University a wonderful experience.

First of all, I would like to thank my thesis supervisors, Drs. G. Fisher and D. Willson, for their excellent guidance and strong support during my graduate studies. They encouraged me to work on the tests for stochastic dominance when I initially proposed the idea of the tests in the late 1992. This thesis is a product of the effort along that path. Their feedbacks and comments have undoubtedly stimulated my thinking and have been instrumental in the research program.

Thanks also go to many professors at Concordia University. They are Drs. S. Ahsan, A. Anastasopoulos, C. Belzil, B. Campbell, A. Hochstein, I. Irvine, V. H. Kirpalani, L. Kryzanowski, Z. R. Liu, J. McIntosh, M. Sampson, F. Sbrocchi, E. Siggel, W. A. Sims, M. Singer, J. Thomberlin, and S. Wang. They either taught and/or supported me during my career development and graduate studies. I must also thank C.M. Beach (Queen's University), P. Callier (World Bank), D. Hester (University of Wisconsin-Madison), D. Huang (People's University), M. Hung (McGill University), G. A. Whitmore (McGill University), S. Wu (Beijing University), Z. Xu (People's University), and X. Yu (People's University) for their support.

I would give special thanks to my fellow graduate students at Concordia University. Several of them kindly offered their comments on my research. They are E. Bristow (Economics) , B. Yang (Mathematics and Statistics), and X. Zhang (Economics).

During my studies at Concordia University, the University and Department awarded me scholarships, grants from the professional development fund, and teaching fellowships. Without such generous financial assistance, I would not be able to accomplish so much.

Finally, I must give thanks to my wife, Xiaoli, and my son, Jia, for their understanding and companionship over the last five years.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Stochastic dominance analysis has widespread applications in various areas of economics and finance. For example, it is used for the evaluation of income distributions of an economy over time or across economies. It is also used for ranking risky prospects in financial economics. The theory of stochastic dominance analysis was developed by Hadar and Russell (1969), Hanoch and Levy (1969), Rothschild and Stiglitz (1970) and Whitmore (1970). Later, numerous papers on or using stochastic dominance analysis have been published, in particular in welfare economics and finance.

The concepts of stochastic dominance are theoretically useful because they allow for the ranking of income distribution or financial asset returns using criteria that are generally acceptable to the majority of researchers without making restrictions about the ranking of preferences. Empirical tests of a dominance relationship provide important information concerning real-world income distributions and financial asset returns.

The applications of stochastic dominance concepts—first-, second-, and third-degree stochastic dominance (denoted as FSD, SSD, and TSD, respectively)— require comparing either the distribution functions or the quantile functions of random variables. In the early stage of this literature, the procedures used for stochastic dominance analysis were typically algorithms rather than statistical tests [See Levy and Hanoch (1970), Porter, Wart and Ferguson (1973) and Vickson (1977), Vickson and Altman (1977), Levy and Kroll (1979), and Kroll and Levy (1980)]. In other words, these procedures generally ignored sampling errors and issues of statistical inference. For example, Tolley and Pope (1988) point out, "In spite of the numerous applications of stochastic dominance techniques, sampling errors are seldom considered."

More recent authors have addressed this issue, although the approaches used have been noticeably restrictive. Some tests are restricted to classes of known parametric distributions [see, for example, Deshpande and Singh (1985), and Stein and Pfaffenberger (1986)]. Since dominance criteria are attractive primarily because they allow for a ranking of returns on risky assets or income distributions, while placing only weak restrictions on preferences, it is important that dominance tests retain a degree of generality, and hence remain nonparametric in nature.

In addition, existing dominance tests often specify the null hypothesis under consideration improperly [see, for example, Tolley and Pope (1988), Bishop, Formby, and Thistle (1989), Bishop, Chakraborti, and Thistle (1989), and McFadden (1989)]. For example, most test procedures make use of the null hypothesis under which two distribution or quantile functions are identical. The hypothesis of dominance may be viewed as an hypothesis of inequality in a particular direction between two distributions. If such an hypothesis is rejected, then dominance cannot be sustained, a result that may or may not be caused by the equality of the two distributions. On the other hand, if the null hypothesis of equality is rejected, then the two distribution cannot be said to be equal, but the cause may or may not be that one dominates the the other. Thus the null hypothesis of equality is not very helpful in providing information about dominance [see Levy (1992), p.574].

Finally, existing test procedures are not appropriate for data that exhibit weak dependence within samples and/or association between samples. This is particularly important in finance applications, where there is substantial evidence suggesting that returns on risky assets are not *i.i.d.*; instead, the emerging empirical consensus suggests that while returns on assets may be unconditionally homogeneous (i.e. identically distributed), they are conditionally heterogeneous with time varying conditional variances (and possibly higher-order moments as well). Allowing for association between the random variables under consideration is also desirable, because returns on assets are obviously determined jointly, and a positive correlation between returns will reduce the variance of the differences between distribution or quantile function estimates.

This thesis develops new distribution-free tests for FSD and SSD, which avoid the limitations of the existing test procedures. The rest of the thesis is organized as follows. Chapter 2 reviews the basic concepts of stochastic dominance. Chapter 3 reviews the literature on the existing test procedures for FSD and SSD and identifies their weaknesses. Chapter 4 develops new distribution-free tests and statistics for FSD and SSD.[1] In Chapter 5, the finite sample properties of the proposed test procedure are evaluated using Monte Carlo simulations. In Chapter 6, an application is demonstrated. Finally, summary and discussion are offered in Chapter 7.

[1]These two types of stochastic dominance relations are mostly used in practical work. The author is working on the third-degree stochastic dominance testing as another project.

Chapter 2

Basic Concepts of Stochastic Dominance

This chapter reviews briefly the basic concepts of stochastic dominance. It serves to provide a background for later discussion. In particular, the basic characteristic of the dominance relations will be highlighted.

2.1 Utility or Social Welfare Functions

Scientific decision-making is commonly based on a preference ordering of available choices. When these choices are uncertain and can be characterized by random variables defined on a probability space, economists typically use the expected utility function to order choices. To be more specific, assume that two risky prospects or random variables X and Y have a joint probability distribution F_{XY}) with well defined support. The marginal distribution of X is F_X, and that of Y is F_Y. Under the expected utility hypothesis, X is said to be preferred to Y if

$$E_{F_X}[U(X)] \geq E_{F_Y}[U(Y)] \tag{2.1}$$

or

$$\int_a^b U(t)dF_X(t) \geq \int_a^b U(t)dF_Y(t), \tag{2.2}$$

where $U(.)$ is a utility function, $E_{F_X}(.)$ and $E_{F_Y}(.)$ are the mathematical expectations under F_X and F_Y, respectively, and $[a,b]$ is the support of the random variable X and Y.[1]

Obviously, two distribution functions and one utility function must be specified in the above evaluation equation. While the distribution functions could be estimated from historical data, the selection of a utility function is somewhat arbitrary and must be done with great care. When the chosen parametric utility function is not satisfactory, then the expected utility evaluation also becomes unsatisfactory.

Stochastic dominance may be applied when only general characteristics of the underlying utility function are known. There is thus no need to impose a specific parametric form of utility function, as would be required in (2.1) and (2.2). In general, three classes of utility functions are commonly used.

Definition 1 *Denote three classes of utility functions as U_i for $i = 1,2,3$. U_1 includes all the functions u with $u' \geq 0$; U_2 includes all the functions u with $u' \geq 0$ and $u'' \leq 0$; and U_3 includes all the functions u with $u' \geq 0$, $u'' \leq 0$ and $u''' \geq 0$.*

U_1 is a class of monotonically increasing utility functions. This class of functions characterizes a preference for higher utility. U_2 is a class of monotonically increasing and concave utility functions. This class of functions characterizes a preference for higher utility and lower risk. U_3 is a more restrictive class of utility functions. This class, together with third-degree stochastic dominance, are only briefly introduced. The test procedures developed in this paper are mainly for first- and second-degree stochastic dominance.

---

[1]In a more general situation, $a = -\infty$ and $b = +\infty$.

2.2 Decision Rules Based on Distribution Functions

One set of decision rules for stochastic dominance relations is based on distribution functions.

Definition 2 *X dominates Y in the first-degree, second-degree, and third-degree, denoted by XD_1Y, XD_2Y, and XD_3Y, respectively, if*

$$F_Y(w) - F_X(w) \geq 0 \ \forall w \in [a,b] \ (XD_1Y); \tag{2.3}$$

$$\int_a^w [F_Y(t) - F_X(t)]dt \geq 0 \ \forall w \in [a,b] \ (XD_2Y); \tag{2.4}$$

and

$$\int_a^v \int_a^w [F_Y(t) - F_X(t)]dtdw \geq 0 \ \forall w, v \in [a,b], \ \text{and} \ E_{F_X}(X) \geq E_{F_Y}(Y) \ (XD_3Y). \tag{2.5}$$

Alternatively, strict dominance in the first-degree, second-degree, and third-degree are defined by (2.3), (2.4) and (2.5) with at least one strict inequality.

Theorem 1 gives a one-to-one correspondence relation between stochastic dominance in the i-th-degree and a preference ordering based on the expected utility functions of the i-th-class ($i = 1,2,3$).

Theorem 1 *Stochastic dominance in the i-th-degree, $i = 1,2,3$, (FSD, SSD, and TSD) and a preference ordering based on the expected utility functions of the i-th-class, $E(U_i)$, $i = 1,2,3$, are related as follows:*

$$F_Y(w) - F_X(w) \geq 0 \ \forall w \in [a,b] \Longleftrightarrow E_{F_X}u(X) \geq E_{F_Y}u(Y) \ \forall u \in U_1; \tag{2.6}$$

$$\int_a^w [F_Y(t) - F_X(t)]dt \geq 0 \ \forall w \in [a,b] \Longleftrightarrow E_{F_X}u(X) \geq E_{F_Y}u(Y) \ \forall u \in U_2; \tag{2.7}$$

$$\int_a^v \int_a^w [F_Y(t) - F_X(t)]dtdw \geq 0 \ \forall w, v \in [a,b] \ \ and \ \ E_{F_X}(X) \geq E_{F_Y}(Y) \qquad (2.8)$$

$$\iff E_{F_X}u(X) \geq E_{F_Y}u(Y) \ \forall u \in U_3.$$

Proof. See Hadar and Russell (1969), Hanoch and Levy (1969) and Rothschild and Stiglitz (1970) for the proofs of FSD and SSD. For the proof of TSD, see Whitmore (1970). See also Bawa (1975). □

It is clear that FSD is the strongest form stochastic dominance in the sense that it implies SSD, while SSD implies TSD, i.e., FSD ⇒ SSD ⇒ TSD.

Figures 2.1 and 2.2 illustrate XD_1Y and XD_2Y using the distribution functions F_X and F_Y. In Figure 2.1, F_X is to the right of F_Y, i.e., F_X is less than or equal to F_Y for all $w \in [a,b]$. In Figure 2.2, the sum of the areas created by the two curves from the point of $w = a$ to any $w \in (a,b]$ is always greater than or equal to zero, i.e., $\int_a^w [F_Y(t) - F_X(t)]dt \geq 0$ for all $w \in [a,b]$.

2.3 Decision Rules Based on Quantile Functions

Decision rules for stochastic dominance relations can also be based on quantile functions.

Definition 3 *If the distribution F is strictly monotonic, the quantile function of order p, $Q(p)$, is the inverse function of the distribution function F. $Q(p)$ is given by $Pr[w \leq Q(p)] = p$. If the distribution function is weakly monotonic, then $Q(p) = \inf\{w : F(w) \geq p, 0 \leq p \leq 1\}$.*

Theorem 2 corresponds to Theorem 1 but is put in terms of quantile functions.

Theorem 2 *X dominates Y in the first-degree, second-degree, and third-degree, denoted by XD_1Y, XD_2Y, and XD_3Y, respectively, if*

$$Q_X(p) - Q_Y(p) \geq 0 \ \forall p \in [0,1] \ (XD_1Y); \tag{2.9}$$

$$\int_0^p [Q_X(t) - Q_Y(t)]dt \geq 0 \ \forall p \in [0,1] \ (XD_2Y); \tag{2.10}$$

$$\int_0^p \int_0^z [Q_X(t) - Q_Y(t)]dtdz \geq 0 \ \forall p, z \in [0,1] \ (XD_3Y), \tag{2.11}$$

and

$$\int_0^1 [Q_X(t) - Q_Y(t)]dt \geq 0.$$

Alternatively, strict dominance in the first-degree, second-degree, and third-degree are defined by (2.9), (2.10) and (2.11) with at least one strict inequality.

Proof. See Levy and Kroll (1978). □

Figures 2.3 and 2.4 show XD_1Y and XD_2Y using the quantiles Q_X and Q_Y. In Figure 2.3, $Q_X(p)$ is to the left of $Q_Y(p)$, i.e., $Q_X(p) - Q_Y(p) \geq 0$ for all $p \in [0,1]$. In Figure 2.4, the sum of the areas surrounded by the two curves from the point of $p = 0$ to any $p \in (0,1]$ is always greater than or equal to zero, i.e., $\int_0^p [Q_X(t) - Q_Y(t)]dt \geq 0$ for $p \in [0,1]$.

While dominance criteria and tests of dominance relationships can be based on either distribution or quantile functions, and their associated empirical estimates, the tests developed here will be based on sample quantile estimates, although this is primarily a matter of convenience. Tests based on sample distribution function estimates (i.e. sample proportions) may be even simpler to compute under more restrictive conditions than those considered in this thesis [see, for example, Anderson (1994)].[2] However, under the general assumptions considered here, the tests are roughly equal in computational complexity; the tests based on sample quantiles are

[2] In the *i.i.d.* case with two independent samples, the variance-covariance matrix of the differences in sample proportions is quite simple to compute, while the test based on sample quantiles requires nonparametric density estimation for the variance-covariance matrix of the estimates.

conceptually simpler, because the domain of the quantile functions for both random variables being compared is the interval $[0,1]$, while the domain of the distribution functions are case-specific and may not be the same, which increases the possibility of 'sampling zeros' in tests based on sample proportions.

Applications of stochastic dominance are important in several areas in economics and finance. For example, in financial economics, Theorems 1 and 2 can be applied directly to an analysis of asset choice. Another example is the income distribution literature. Here the terminology used is a little different from that in financial economics. Consider generalized Lorenz (GL) dominance as given in Shorrocks (1983) and let F_X and F_Y denote two income distributions. GL curves can be defined over the corresponding quantile functions Q_X and Q_Y, i.e., $GL_X(p) = \int_0^p Q_X(t)dt$ and $GL_Y(p) = \int_0^p Q_Y(t)dt$. GL (*strict*) dominance of an income distribution F_X over another income distribution F_Y is defined as $GL_X(p) \geq GL_Y(p)\ \forall p \in [0,1]$ (*with a strict inequality for a least one p*). This is equivalent to SSD in Theorem 2.[3]

2.4 Basic Characteristic of Dominance Relations

Since this thesis focuses only on FSD and SSD in terms of quantile functions, the basic characteristic of dominance relations given in FSD and SSD will be reviewed in the context of quantile functions or transformations of quantile functions.

Empirically, a researcher must decide whether to treat the sample quantile functions as stochastic processes or to estimate the functions at points. In the former case, asymptotic theory has not been developed for these types of tests, although it certainly is possible. For the latter, asymptotic theory is available, but the number and locations of points must be determined by a researcher.

Dominance relationship can be defined in either regular (weak) or strict form. It turns out as shown later that the regular dominance relationship (dominance rela-

[3]Shorrocks (1983) shows that the GL dominance is equivalent to preference by all increasing, anonymous, equality-preferring social welfare functions.

tionship, hereafter) is convenient for developing test statistics. When this dominance relationship is employed, the tests aim at providing statistical evidence that the weak dominance relationship does not exist between the two distribution or quantile functions being compared. Because the two objects being compared are interchangeable, the direction in which a dominance relation is evaluated is important in providing insight into the relation.

From Theorem 2, it is clear that the dominance relations in the sense of FSD and SSD can not be simply explained as the equality of two quantile functions (in the case of FSD) or of two transformations of quantile functions (in the case of SSD). In other words, testing for stochastic dominance relations is not equivalent to testing for the equality of two functional forms.

To explain the above point, let θ_X and θ_Y be either quantile functions (in the case of FSD) or transformations of quantile functions (in the case of SSD) of X and Y, respectively. The tests for FSD and SSD should be able to differentiate the following two cases statistically, viz., H_0: $\theta_Y - \theta_X \geq 0$ (dominance) against H_a: $\theta_Y - \theta_X \not\geq 0$ (non-dominance). These hypotheses are very different from another set of hypotheses. H_0: $\theta_Y - \theta_X = 0$ (equality) against H_a: $\theta_Y - \theta_X \neq 0$ (inequality). The former is a case where dominance is against non-dominance; while the latter is a case where equality is against inequality. Existing test procedures generally misspecify the relations to be tested. More specifically, those procedures are designed to test the latter while they are claimed to be useful to test the former.

Figure 2.1: The stochastic dominance: XD_1Y

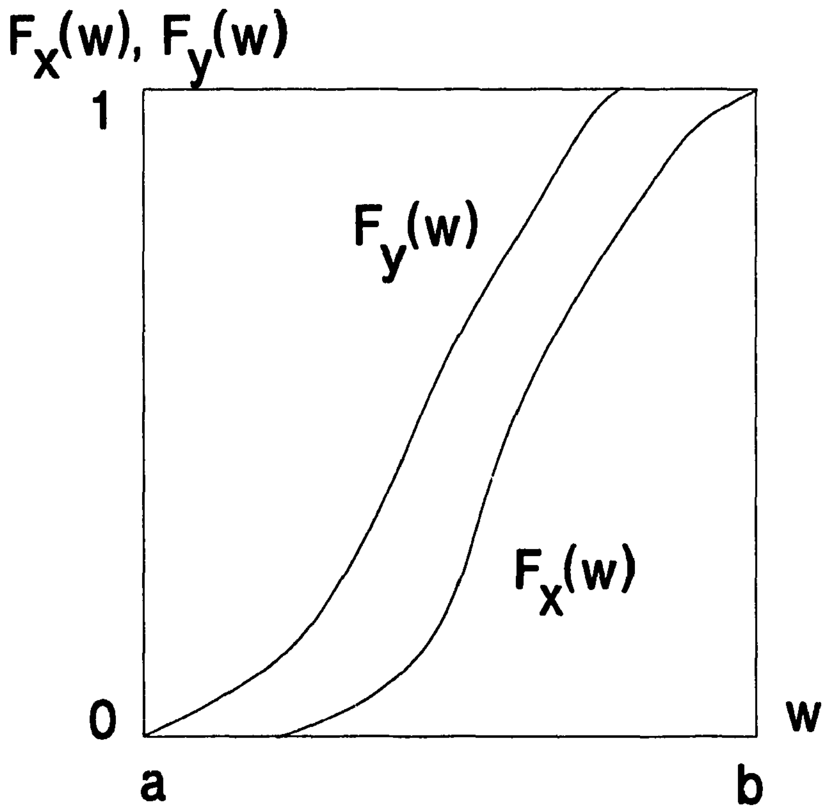


Figure 2.2: The stochastic dominance: $X D_2 Y$

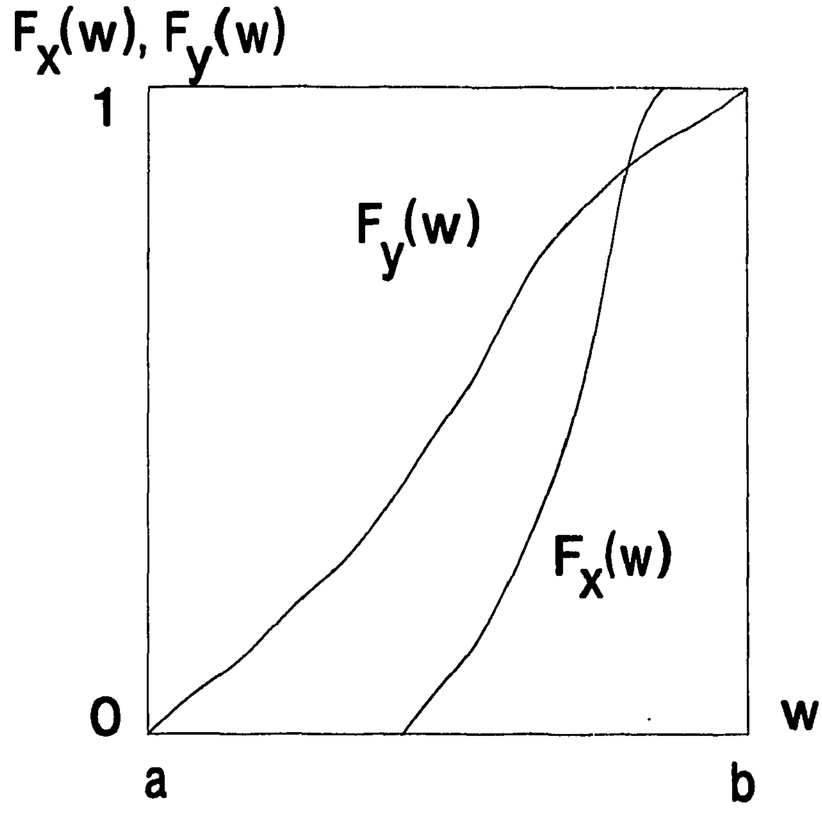


Figure 2.3: The quantile condition of stochastic dominance: XD_1Y

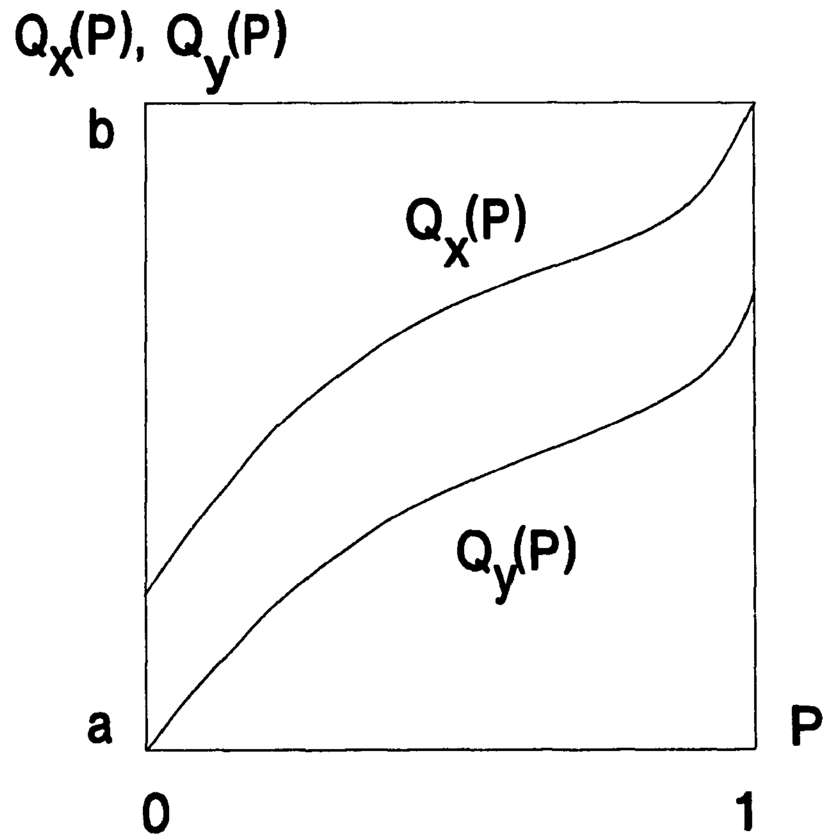
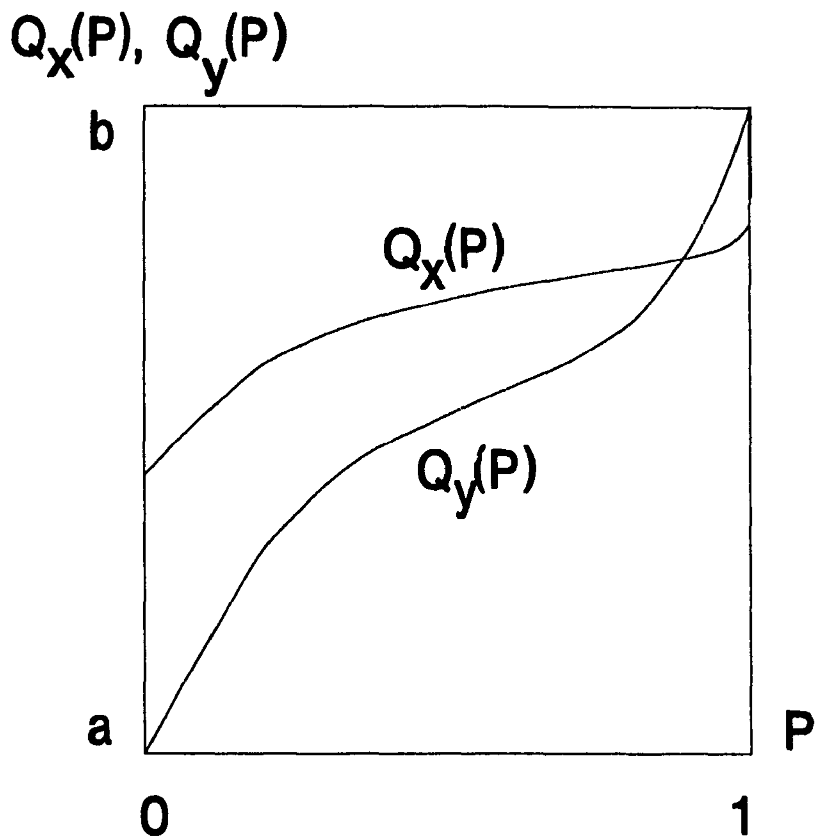


Figure 2.4: The quantile condition of stochastic dominance: XD_2Y



Chapter 3

Literature Review

This chapter reviews the literature of existing algorithms and test procedures for stochastic dominance analysis. These are the: *early algorithm* [Levy and Hanoch (1970) and Porter, Wart and Ferguson (1973)], *conventional quantile approach* [Levy and Kroll (1979)], *parametric test procedure* [Whitmore and Findlay (1978), Deshpande and Singh (1985), and Stein and Pfaffenberger (1986)], *randomization test procedure* [Tolley and Pope (1988)], *SMM multiple-comparison test procedure* [Richmond (1982), Beach and Richmond (1985), and Bishop, Formby, and Thistle (1989)], *Smirnov-type distribution-free test procedure* [Whitmore and Findlay (1978), and McFadden (1989)], and *Wald-type distribution-free test procedure* [Bishop, Chakraborti, and Thistle (1989)].

The early algorithm and conventional quantile approach are the procedures that are not statistical procedures. The parametric test procedure is a statistical procedure but is restricted to the family of parametric distribution functions. The randomization test procedure, SMM multiple-comparison test procedure, Smirnov-type distribution-free test procedure, and Wald-type distribution-free test procedure vary in their constructs, but they are actually useful for testing the equality of two (distribution/quantile or transformations of distribution/quantile) functions, instead of the dominance relation of the two. Further, these test procedures are not designed to accommodate complex data structures such as weak dependence within samples and/or association between samples. The following review will highlight the deficiencies of the existing procedures, and demonstrate the need for new tests that avoid the deficiencies.

Let X and Y be two random variables with the marginal distribution functions F_X and F_Y, respectively. Denote the samples from X and Y by $\{x_t\}_{t=1}^{T_X}$ and $\{y_t\}_{t=1}^{T_Y}$, respectively. T_X and T_Y are the numbers of observations in the samples from X and Y. Sometimes, T is used if $T = T_X = T_Y$. Observations in each sample can be arranged in increasing order, i.e., $x_{(1)} \leq x_{(2)} \cdots \leq x_{(T)}$ and $y_{(1)} \leq y_{(2)} \cdots \leq y_{(T)}$, where $x_{(t)}$ and $y_{(t)}$ are t-th order statistics of X and Y. The empirical marginal distribution functions are denoted as F_{XT} and F_{YT}. Occasionally, F_0 is used to denote a known parametric distribution for a random variable X_0.

3.1 Early Algorithm

The *early algorithm* for empirical FSD and SSD analysis relies on some useful short-cuts. It implicitly uses the salient features of the stochastic dominance criteria without actually examining distribution functions directly. These short-cuts [1] are as follows:

1. One risky prospect dominates another only if it has a mean equal to or greater than the mean of the other.

2. If two risky prospects have the same mean, the one with the greater variance cannot dominate the other.

3. If the lowest observation of one risky prospect is below the lowest observation of a second, the first cannot dominate the second.

[1] These were summarised by Kroll and Levy (1980).

4. Dominance of one risky prospect over another by FSD implies dominance by SSD, and dominance by SSD implies dominance by TSD.

5. If one risky prospect does not dominate another in terms of TSD, then there is also no dominance by SSD, which in turn implies no dominance by FSD.

6. If a risky prospect is dominated by at least one risky prospect, it can be excluded from the efficient set and there is no need for further comparisons between it and other risky prospects.

The early algorithm is merely an algorithm. It can best be interpreted as a set of decision rules in a world of certainty. However, the stochastic dominance is about the decision rules in a world of uncertainty. Thus, the early algorithm is undoubtedly inappropriate.

3.2 Conventional Quantile Approach

The *conventional quantile approach* is an ordering algorithm rather than a statistical test procedure. The basic idea of this approach is to compute the quantiles of two samples and compare them. If two risky prospects, X and Y, are evaluated in the sense of FSD and SSD, the following decision rules are used:

1. XD_1Y if and only if $x_{(t)} \geq y_{(t)} \forall\, t$ with at least one strict inequality.

2. XD_2Y if and only if $\sum_{t=1}^{j} x_{(t)} \geq \sum_{t=1}^{j} y_{(t)} \forall\, j$ with at least one strict inequality.

This approach is well-known and widely used. However, it ignores sampling errors, i.e., it treats a sample distribution function as the corresponding population distribution function, and hence it does not provide any basis for statistical inference. Stein and Pfeffenberger (1986) use the specific parametric distributions, such as the normal and lognormal distributions, to show the error probability associated with this approach for a wide range of parameter values.

3.3 Parametric Test Procedure

The *parametric test procedure* was proposed to deal with the problems in the early algorithm and conventional quantile approaches. Unfortunately, the tests basically compare either parameters of distribution functions or distribution functions within a small family of parametric distribution functions. Two variants of parametric tests are briefly reviewed here.

The first example is a parameter test based on parametric distribution functions due to Stein and Pfeffenberger (1986). In the following, the normal distribution is used as an example to demonstrate how the test is constructed. X has mean μ_X and variance σ_X^2, while Y has mean μ_Y and variance σ_Y^2. To test dominance relations, first transform X to a standard normal random variable with standardized distribution function F_X^s; then compare the standard normal distribution F_X^s to a normal distribution F_Y^s with parameters, $\mu = (\mu_Y - \mu_X)/\sigma_X$ and $\sigma = \sigma_Y/\sigma_X$. The decision rules for FSD are: XD_1Y if $\mu < 0$ and $\sigma = 1$; YD_1X if $\mu > 0$ and $\sigma = 1$; and there is no dominance otherwise. The decision rules for SSD are: XD_2Y if $\mu \leq 0$ and $\sigma \geq 1$; YD_2X if $\mu \geq 0$ and $\sigma \leq 1$; and there is no dominance otherwise.[2]

The second example is the test for SSD due to Deshpande and Singh (1985). The idea of the test is based on a claim that comparing a random prospect X, which has the distribution function F_X, with another random prospect X_0 with a known parametric distribution function F_0, is equivalent to testing for SSD in the case of utility functions belonging to U_2. The authors propose the following null and alternative hypotheses: $H_0 : F_X = F_0$ against $H_a : F_X D_2 F_0$. It is assumed that distribution functions, F_X and F_0, are continuous and that their first two moments exist. Let $d_{F_X,F_0}(w) = \int_a^w (F_X(t) - F_0(t))dt$, where $w \in [a, b]$, and $D_{F_X,F_0} = \int_a^b d_{F_X,F_0} dF_0(w)$. D_{F_X,F_0} is a measure of distance between F_X and F_0. According to Deshpande and Singh, $D_{F_X,F_0} = 0$ under H_0 and $D_{F_X,F_0} < 0$ under H_a. Let F_{XT} be the corre-

[2]See Stein and Pfeffenberger (1986) for details.

sponding empirical distribution function.[3] Define $d_T(w) = \int_a^w (F_{XT}(t) - F_0(t))dt$ and $D_T = \int_a^b d_T(w)dF_0(w)$. D_T is an estimator of D_{F_X,F_0}. Under H_0, D_T will take values near zero whereas under H_a it will be negative and has a large absolute value. The authors suggest that the decision rules are: if $D_T < C_\alpha$ (where C_α is an appropriately chosen critical point so that the test has the desired size α), reject H_0; otherwise, do not reject H_0. The two authors demonstrate that if

1. a random sample, $\{x_t\}_{t=1}^T$, is drawn from the distribution F_X;

2. F_0 is a known parametric distribution;

3. $D_T = \frac{1}{T}\sum_{t=1}^T y_t - A$, where $y_t = \int_{x_t}^b (w - x_t)dF_0(w)$, $t = 1, 2, \cdots, T$ and $A = \int_a^b F_0(w)(1 - F_0(w))dw$; and

4. $S_T = \frac{D_T - E_{F_X}(D_T)}{\sigma_{F_X}(D_T)}$, where $E_{F_X}(D_T)$ and $\sigma_{F_X}(D_T)$ denote the mean and standard deviation of D_T with respect to the distribution function F_X;

then S_T converges in distribution to a standard normal distribution.

Obviously, the above test is only useful if F_0 has a known parametric form. Furthermore, the test has two other problems. First, the derivation of the test-statistic relies on an erroneous claim, viz., $D_{F_X,F_0} < 0 \Rightarrow d_{F_X,F_0}(w) < 0 \; \forall \, w \in [a, b]$. Obviously, this is not necessarily true. Second, the test procedure has the null hypothesis under which the two distribution functions are identical. This setup is not desirable for evaluating dominance relationship. The second problem is reflected in several test procedures for SSD.

[3] Many possible estimators for a distribution function are available in the literature. However, Deshpande and Singh (1985) do not give the specification of F_{XT}.

3.4 Randomization Test Procedure

The *randomization test procedure*[4] was proposed for stochastic dominance analysis by Tolley and Pope (1988). Although this test can be applied for both FSD and SSD, Tolley and Pope's work focuses on a test for SSD. The hypotheses for XD_2Y can be established as $H_0 : X\not{D}_2Y$ against $H_a : XD_2Y$. Let the sample counterpart of $D(w) = \int_a^w (F_Y(t) - F_X(t))dt$ be $D^*(w) = w\left(\frac{k}{T_X} - \frac{l}{T_Y}\right) + \frac{1}{T_Y}\sum_{t=1}^{l} y_{(t)} - \frac{1}{T_X}\sum_{t=1}^{k} x_{(t)}$, where $w \in [a,b]$; $x_{(t)}$ and $y_{(t)}$ are the ordered observations of $\{x_t\}_{t=1}^{T_X}$ and $\{y_t\}_{t=1}^{T_X}$, respectively; k is such that $x_{(k)} \leq w < x_{(k+1)}$; and l is such that $y_{(l)} \leq w < y_{(l+1)}$.[5] The critical regions, C_α and C'_α corresponding to the significance level α for the statistic $D^*(w)$ is formulated from the following logical steps for testing H_0 against H_a:

1. If $D^*(w)$ is significantly less than zero for any value of w, then H_0 is not rejected. If $\min_w D^*(w) < C_\alpha < 0$ for some value of C_α, then the test procedure is stopped and H_0 is not rejected in favor of H_a.

2. If $D^*(w)$ is not significantly less than zero for any w and if $D^*(w)$ is significantly greater than zero for some w, then H_0 is not rejected in favor of H_a. Thus, given that $\min_w D^*(w) \not< C_\alpha$, if $\max_w D^*(w) > C'_\alpha > 0$ for some fixed C'_α, then H_0 is rejected in favor of H_a.

The remaining work of the test procedure is to determine values of C_α and C'_α. These are determined empirically using a permutation test procedure on the desired significance level α.[6]

Although H_0 and H_a are formulated correctly, the Tolley and Pope test procedure in fact uses an implicit assumption under the null hypothesis that the two distributions are identical [i.e., no treatment effects. See Tolley and Pope (1988),

[4]The randomization tests do not require samples to be random. While this is an advantage, the tests are not appropriate for statistical inference of the population. See Noreen (1989).

[5]For the derivation, see Appendix of Tolley and Pope (1988).

[6]See Tolley and Pope (1988) for details.

p.696 and Edington (1980), p.11.]. Once again, this test procedure is designed to test the equality of two distribution functions against a violation of such equality. rather than the dominance relationship.

3.5 SMM Multiple-Comparison Test Procedure

The multiple-comparison or simultaneous inference procedure is based on a technique developed by Richmond (1982) for constructing joint confidence intervals for all possible linear combinations of means of multivariate normal distributions. Beach and Richmond (1985) apply this technique to construct joint confidence intervals for Lorenz ordinates. Based on the Richmond's results, the SMM multiple-comparison tests were developed [See Bishop, Formby, and Thistle (1989)].

Richmond's joint confidence interval procedure uses the studentized maximum modulus (SMM) distribution. The basic idea is given below. Assume $\mathbf{x} = (x_1, x_2, \ldots x_K)'$ is a K-dimensional random vector distributed $N(\mu, \sigma^2\Sigma)$. Let s^2 be an estimator of σ^2, where $vs^2/\sigma^2 \sim \chi^2(v)$ independently of $\bar{\mathbf{x}}$. Then, if Σ is a diagonal matrix with diagonal elements d_{ii}, dividing by the standard errors, $\Sigma^{-1/2}(\mathbf{x} - \mu)/s$ has a limiting multivariate standard normal distribution. The largest of these "studentized" random variables, $\widetilde{m} = \max_i\{|x_i - \mu_i|/s\sqrt{d_{ii}}\}$, has the SMM distribution with K and v degrees of freedom, or $\widetilde{m} \sim SMM(K, v)$. If $m_\alpha(K, v)$ is used to denote the upper $100(1-\alpha)$ of the $SMM(K, v)$ distribution, then

$$\Pr[\widetilde{m} \geq m_\alpha(K, v)] = \alpha. \tag{3.1}$$

When Σ is not diagonal, the x_i's are not independent. The above equation becomes an inequality as shown in Hochberg (1974):

$$\Pr[\widetilde{m} \geq m_\alpha(K, v)] \leq \alpha. \tag{3.2}$$

Richmond extends this inequality to linear combinations of the x_i's. Let $\mathbf{B} = [\mathbf{b}_1 \cdots \mathbf{b}_q]'$ be a matrix which contains q linearly independent known K-vectors, where $q \leq K$. Thus, $\mathbf{Bx} \sim N(\mathbf{B}\mu, \sigma^2 \mathbf{B}\Sigma\mathbf{B}')$. Richmond showed that

$$\Pr[|\mathbf{b}_i(\mathbf{x} - \mu)|/s(\mathbf{b}_i'\Sigma\mathbf{b}_i)^{1/2} \leq m_\alpha(q, v), i = 1, 2, \ldots q] \geq 1 - \alpha. \tag{3.3}$$

This inequality give the joint $100(1-\alpha)$ confidence intervals for all possible linear combinations of the means. This inequality is also used to derive the multiple-comparison tests.

As an example, consider the multiple-comparison test for the differences in the conditional means of two samples. Suppose $\widehat{\mathbf{M}}_X$ and $\widehat{\mathbf{M}}_Y$ are estimators of the vectors of conditional means from two independent samples of T_X and T_Y. Under the null hypothesis, the two samples are from the *same* distribution, i.e., H_0: $M_X = M_Y$. Let $\mathbf{Z} = [Z_1 \cdots Z_K]'$, where Z_j is the j-the difference in the conditional means of two samples that are, it is assumed, from the *same* distribution. The vector $\mathbf{Z}$ has an asymptotic distribution $N(\mathbf{0}, \mathbf{I})$. The Hochberg-Richmond inequality implies

$$\Pr[|Z_j| \leq m_\alpha(K, \infty), j = 1, 2, \ldots K] \geq 1 - \alpha. \tag{3.4}$$

Multiple comparisons of pairs of conditional means can be made using the statistics Z_j. To carry out the multiple comparisons, the Z_j's are tested as $SMM(K, \infty)$ variates. Tables of the SMM distribution are given in Stoline and Ury (1979).

The set of K multiple-comparison tests implies a joint test for equality of the vectors of conditional means. This implied joint test is a union-intersection test. The implied joint test does not reject the overall null hypothesis H_0: $M_X = M_Y$ if and only if all of the component null hypotheses are not rejected. Formally, the test of the j-th component hypothesis is : Reject H_0^j if and only if $|Z_j| \geq m_\alpha(K, \infty)$; otherwise do not reject H_0^j. The test of the overall hypothesis is: Reject H_0 if and only if H_0^j is rejected for any j; otherwise do not reject H_0. Therefore, the rejection (acceptance) region of the implied joint test is the union (intersection) of the rejection (acceptance) regions of the individual component tests.

It is clear that the multiple-comparison tests are designed to test the equality of two vectors of parameters against the violation of such equality. These tests for the equality relation are not completely suitable for testing stochastic dominance relationships.

3.6 Smirnov-Type Distribution-Free Test Procedure

McFadden (1989) developed a *Smirnov-type distribution-free test procedure* for stochastic dominance. A brief summary of his tests for FSD and SSD is given as below. The tests are suitable to the data generating processes where X and Y are statistically independent and the samples from X and Y are *i.i.d.*. It is assumed for simplicity that $T = T_X = T_Y$. The distribution functions F_X and F_Y are not restricted to any parametric form. Define F_{XT} to be the empirical distribution function based on the sample observations from X, i.e.,

$$F_{XT}(w) = \frac{1}{T}\sum_{t=1}^{T} I(x_t \leq w), \tag{3.5}$$

where $I(A)$ is the indicator function: one if A is true; zero otherwise. $F_{YT}(w)$ is defined as the empirical distribution for Y. To test for XD_1Y, consider $H_0 : F_Y(w) - F_X(w) \geq 0 \ \forall\, w \in [a,b]$ against $H_a : F_Y(w) - F_X(w) < 0$ for some $w \in [a,b]$. For the null hypothesis, a test statistic is based on

$$D_T^* = \max_{w\in[a,b]} D_T(w), \tag{3.6}$$

where $D_T(w) = \sqrt{T}[F_{YT}(w) - F_{XT}(w)]$.

Let $\{z_t\}_{t=1}^{2T}$ denote the *ordered pooled* observations from the X and Y samples. Define d_t to be the indicator that is $+1$ if z_t is from the Y sample, and -1 if z_t is from the X sample. Let $H_{2T}(w)$ denote the empirical distribution function from the pooled observations, and note that $2TH_{2T}(w) = t$ implies $z_t \leq w < z_{t+1}$, where $w \in [a,b]$. Define $D_{Tt} = \frac{1}{\sqrt{T}} \sum_{j=1}^{t} d_j$. Then

$$D_T(w) = \frac{1}{\sqrt{T}}\sum_{j=1}^{2T} d_j I(z_j \leq w) = D_{Tt} \tag{3.7}$$

implies

$$D_T^* = \max_{1 \le t \le 2T} D_{Tt}. \tag{3.8}$$

This statistic is easy to compute using $D_{T,t+1} = D_{Tt} + \frac{d_{t+1}}{\sqrt{T}}$.[7] Under H_0, the exact distribution of this statistic is

$$P(D_T^* > q\sqrt{T}) = \frac{(T!)^2}{(T-k)!(T+k)!} \tag{3.9}$$

where $k > Tq \ge k-1$ (i.e., k is the smallest integer greater than Tq).[8] For large T, this has the limiting distribution

$$P(D_T^* > q\sqrt{T}) \approx e^{-q^2}\left(1 - \frac{q}{3}\sqrt{\frac{2}{T}} + O\left(\frac{1}{T}\right)\right). \tag{3.10}$$

The McFadden distribution-free test for SSD is based on a similar idea. The hypotheses are H_0: $\int_a^w (F_Y(t) - F_X(t))dt \ge 0 \ \forall \ w \in [a,b]$ against H_a: $\int_a^w (F_Y(t) - F_X(t))dt < 0 \ \forall \ w \in [a,b]$. The statistic is based on

$$S_T^* = \max_{w \in [a,b]} S_T(w) \tag{3.11}$$

where $S_T(w) = \sqrt{T}\int_a^w (F_Y(t) - F_X(t))dt$. Using the notation defined above, let

$$S_{Tt} = \frac{1}{\sqrt{T}} \sum_{j=1}^{t} d_j(z_t - z_j). \tag{3.12}$$

This statistic satisfies the recursion $S_{T,t+1} = (z_{t+1} - z_t)D_{Tt} + S_{Tt}$. For $w \in [a,b]$ and $t = 2TH_{2T}(w)$, implying $z_t \le w < z_{t+1}$, one has

$$\begin{aligned} S_T(w) &= \frac{1}{\sqrt{T}} \sum_{j=1}^{t} d_j(w - z_j). \\ &= (w - z_t)D_{Tt} + S_{Tt}. \end{aligned} \tag{3.13}$$

Since $D_{T,2T} = 0, S_T(w)$ is constant for $w \ge z_{2T}$, implying

$$S_T^* = \max_{1 \le t \le 2T} S_{Tt}. \tag{3.14}$$

[7] D_T^* is the well-known Smirnov statistic [see Durbin (1973)].

[8] See Gnedenko and Koroyuk (1961) and Durbin (1985).

This statistic is easy to compute using $S_{T,t+1} = (z_{t+1} - z_t)D_{Tt} + S_{Tt}$. The distribution of the statistic S_T^* does not appear to have a tractable analytical form. McFadden provides a computational method for calculating significance levels.[9]

McFadden's SSD test suffers from three problems. The first is that the SSD test procedure in fact puts the equality of transformations of two distribution functions under the null hypothesis. The second is that the test-statistic S_T^* does not have a tractable analytical distribution. The third problem is that the test cannot accommodate weak dependence within a sample and association between samples.

3.7 Wald-Type Distribution-Free Test Procedure

The *Wald-type distribution-free test procedure*, was proposed by Bishop, Chakraborti and Thistle (1989). They developed the two test-statistics, T_1 and T_2, which are used jointly to test the dominance relations between two cumulative quantile functions in the context of generalized Lorenz curves. Let Ψ_X and Ψ_Y be K-vectors of cumulative quantiles for random variables X and Y respectively. The first test-statistic T_1 is used to test H_0^1: $\Psi_X = \Psi_Y$ against H_a^1: $\Psi_X \neq \Psi_Y$. T_1 has an asymptotical χ^2 distribution with degrees of freedom K and is defined as $T_1 = (\hat{\Psi}_X - \hat{\Psi}_Y)'[(\hat{\Omega}_X/T_X) + (\hat{\Omega}_Y/T_Y)]^{-1}(\hat{\Psi}_X - \hat{\Psi}_Y)$, where $\hat{\Psi}_X$ and $\hat{\Psi}_Y$ represent the estimates of cumulative quantiles for X and Y, respectively; $\hat{\Omega}_X$ and $\hat{\Omega}_Y$ represent the estimates of variance-covariance matrix for $\hat{\Psi}_X$ and $\hat{\Psi}_Y$, respectively; and T_X and T_Y are the numbers of observations in the X and Y samples, respectively. Bishop and Thistle claimed that the second test-statistic T_2 should be used to test H_0^1: $\Psi_X \leq \Psi_Y$ against H_a^1: $\Psi_X > \Psi_Y$. T_2 has an asymptotic standard normal distribution and is defined as $T_2 = \mathbf{1_K}'(\hat{\Psi}_X - \hat{\Psi}_Y)'/(\mathbf{1_K}'[(\hat{\Omega}_X/T_X) + (\hat{\Omega}_Y/T_Y)]^{-1}\mathbf{1_K})^{1/2}$, where $\mathbf{1_K}$ is a $K \times 1$ vector with all elements unity. Their test procedure suffers from one fundamental problem; the test-statistic T_2 has no power to measure the elementwise distance of two vectors, Ψ_X and Ψ_Y. T_2 can test if $\mathbf{1}'(\Psi_X - \Psi_Y) > 0$ but cannot test if

[9] See McFadden (1989) for details.

$(\Psi_X - \Psi_Y) \geq 0$ with a strict inequality for at least one case. Clearly, the two inequality conditions are mathematically different. The relation, $\mathbf{1}'(\Psi_X - \Psi_Y) > 0$, simply states that the sum of the differences is greater than zero. This is not a condition for SSD. The only relation for SSD is $(\Psi_X - \Psi_Y) \geq 0$. The former does not imply the latter, while the later implies the former. A smaller problem is that the two-step procedure will give a significance level which is far different from the significance levels for each step of the tests.

Clearly, the existing procedures suffer from various problems. Most importantly, these test procedures are actually useful for testing the equality of two distribution or quantile functions, instead of a dominance relationship between the distribution or quantile functions. In addition, these test procedures are not designed to accommodate complex data structures. Chapter 4 develops tests for FSD and SSD that correct these weaknesses.

Chapter 4

New Distribution-Free Tests for Stochastic Dominance

4.1 Introduction

As shown previously, the quantile conditions for FSD and SSD are: XD_1Y if and only if $Q_X(p) - Q_Y(p) \geq 0 \ \forall p \in [0,1]$; and XD_2Y if and only if $\int_0^p [Q_X(t) - Q_Y(t)]dt \geq 0 \ \forall p \in [0,1]$. Alternatively, the quantile condition for SSD can be expressed in terms of cumulative quantiles. The cumulative quantile function is given by

$$\Psi(p) = \int_0^p Q(t)dt \tag{4.1}$$

where $p \in [0,1]$. Thus, XD_2Y if and only if $\Psi_X(p) - \Psi_Y(p) \geq 0 \ \forall p \in [0,1]$.

Decision rules for FSD and SSD have a common structure, and both can be expressed in a more general form. Let θ be used to denote either Q or Ψ. The general form of the decision rule is to test if $\theta_Y - \theta_X \geq 0$ or $\theta_X - \theta_Y \geq 0$. Based on this common structure, the new distribution-free tests for FSD and SSD are developed below.

For testing hypotheses such as H_0: $\theta_Y - \theta_X \geq 0$ against H_a: $\theta_Y - \theta_X \not\geq 0$, test-statistics can be designed using both the restricted estimator of $\theta_Y - \theta_X$, $\tilde{\theta}_Y - \tilde{\theta}_X$, and the unrestricted estimator of $\theta_Y - \theta_X$, $\hat{\theta}_Y - \hat{\theta}_X$. The unrestricted estimator $\hat{\theta}_Y - \hat{\theta}_X$ are *free* in a space $\mathcal{U}$, while the restricted estimator $\tilde{\theta}_Y - \tilde{\theta}_X$ must be in the restricted space $\mathcal{R}$, where $\mathcal{R} \subset \mathcal{U}$ and $\mathcal{R} \neq \mathcal{U}$. A dominance relation can be characterized by a relation in the restricted space $\mathcal{R}$ and this relation is specified under the null hypothesis. If both restricted and unrestricted estimators are in the restricted space $\mathcal{R}$, then the null hypothesis should not be rejected. If the unrestricted estimator is not in the restricted space $\mathcal{R}$, and far away from the restricted estimator in the restricted space $\mathcal{R}$, the null hypothesis will be rejected. The tests should therefore be able to measure the distance between the restricted and unrestricted estimators.

In the following sections, the tests under general conditions will be developed, and then two important variants of the tests under relative restrictive conditions will be derived.

4.2 Tests under General Conditions

4.2.1 Assumptions

The test procedure is proposed under very general conditions that allow for weak dependence within samples and association between samples. The motivation for proposing the tests under these general conditions is that the *i.i.d.* and *independence* requirements are not realistic assumptions for complex data structures such as financial data. Many financial time series, such as stock and bond returns, are not *i.i.d.* but identically weakly dependent. Most of them are also correlated cross-sectionally.[1] When observations in each sample are not *i.i.d.* and a pair of samples are not independently selected, the variance-covariance matrix of the difference of two quantile function estimators will change. Thus, there is a need for dominance

[1] For example, Roll and Ross (1980) describe the common contemporaneous comovements of stock returns as their single most important feature

tests which can accommodate both weak dependence within samples and association between samples. In addition, under some additional restrictions, the test-statistics can be simplified substantially and correspond to the proper tests for FSD and SSD under familiar conditions.

Denote the two random variables $\{X_t, Y_t\}$ as the random vector $\{Z_t\}$, where Z_{tj}, $j = X, Y$, or $X_t = Z_{tX}$ and $Y_t = Z_{tY}$. Let $\{Z_t, -\infty < t < \infty\}$ be a stationary ϕ-mixing sequence of random vectors defined on a probability space $(\Omega, \mathcal{A}, \mathcal{P})$. Thus, if $\mathcal{M}_{-\infty}^{k}$ and $\mathcal{M}_{k+T}^{\infty}$ be respectively the σ-fields generated by $\{Z_t, t \leq k\}$ and $\{Z_t, t \geq k+T\}$, and if $E_1 \in \mathcal{M}_{-\infty}^{k}$ and $E_2 \in \mathcal{M}_{k+T}^{\infty}$, then for all $k(-\infty < k < \infty)$ and $T(\geq 1)$,

$$|P(E_2|E_1) - P(E_2)| \leq \phi(T), \quad \phi(T) \geq 0, \tag{4.2}$$

where $1 \geq \phi(1) \geq \phi(2) \cdots$, and $\lim_{T\to\infty} \phi(T) = 0$.

Assumption 1 *The ϕ-mixing sequence satisfies*

$$\sum_{T=1}^{\infty} [\phi(T)]^{1/2} < \infty. \tag{4.3}$$

Z_t has a bivariate density function $f_{XY}(z)$ or $f_{XY}(x,y)$, and distribution function $F_{XY}(z)$ or $F_{XY}(x,y)$, where $z = (x,y) \in \mathbf{R}^2$. The corresponding marginal density and distribution functions for X and Y are written as $f_X(x)$ and $f_Y(y)$, and $F_X(x)$ and $F_Y(y)$, respectively.

Assumption 2 *$F_{XY}(z)$ is strictly monotonic in some neighborhood of ξ in each of its two coordinates and admits of a differentiable continuous density $f_{XY}(z)$, such that*

$$0 < f_{XY}(z) < \infty. \tag{4.4}$$

Assumption 1 essentially requires that the dependence between observations dies out as the (temporal) distance between them increases. Assumption 1 is now standard in empirical finance [see Lo and McKinlay (1988)]. Assumption 2 ensures that the inverse distribution function exists and is well defined. This assumption is important for consistent estimation of the variance and covariance matrix of the point estimates of quantile functions. Assumptions 1 and 2 are required for developing the asymptotic distribution of the sample quantiles given in Lemma 1.

The empirical distribution function for j-th variate is given by

$$F_{Tj}(w) = \frac{1}{n}\sum_{t=1}^{T} I(Z_{tj} \leq w), \tag{4.5}$$

where w is in the support, $j = X, Y$, and $I(A)$ is the indicator function: one if A holds; zero otherwise. K proportions or probabilities and K corresponding quantiles are chosen for both $\{X_t\}$ and $\{Y_t\}$. $\xi = [\xi^X_{(1)}, \xi^X_{(2)}, \cdots \xi^X_{(K)}, \xi^Y_{(1)}, \xi^Y_{(2)}, \cdots \xi^Y_{(K)}]'$ denotes a $2K \times 1$ vector of quantiles defined in $\mathbf{R}^2$. $P\left\{Z_{tj} \leq \xi^j_{(i)}\right\} = p^j_{(i)}$, where $0 < p^j_{(i)} < 1$, and $j = X, Y$. Since it is assumed that $p^X_{(i)} = p^Y_{(i)}$ for all i, $P\left\{Z_{tj} \leq \xi^j_{(i)}\right\} = p^j_{(i)}$ can be written as $P\left\{Z_{tj} \leq \xi^j_{(i)}\right\} = p_i$, i.e., the superscript j of $p^j_{(i)}$ can be suppressed. Thus, the $2K \times 1$ vector, ξ , can be alternatively expressed as

$$\begin{aligned} \xi &= Q_Z(P) \qquad (4.6) \\ &= [Q_X(P), Q_Y(P)]' \\ &= [Q_X(p_1), Q_X(p_2), \cdots Q_X(p_K), Q_Y(p_1), Q_Y(p_2), \cdots Q_Y(p_K)]'. \end{aligned}$$

If we denote the two samples as $\{x_t\}^T_{t=1}$ and $\{y_t\}^T_{t=1}$, observations in each sample can be arranged in increasing order, i.e., $x_{(1)} \leq x_{(2)} \cdots \leq x_{(T)}$ and $y_{(1)} \leq y_{(?)} \cdots \leq y_{(T)}$, where $x_{(t)}$ and $y_{(t)}$ are t-th order statistics of X and Y, respectively. The sample quantiles of order p for X and Y are denoted as $\widehat{Q}_{TX}(p) = x_{([Tp]+1)}$ and $\widehat{Q}_{TY}(p) = y_{([Tp]+1)}$, respectively, where $[Tp]$ refers to the largest integers that are less than or equal to Tp. For a finite set of quantiles, such as $\widehat{Q}_{TX}(P) = \{x_{([Tp_i]+1)} | i = 1, \ldots, K\}$, and $\widehat{Q}_{TY}(P) = \{y_{([Tp_i]+1)} | i = 1, \ldots, K\}$, associated with the the abscissae

$P = \{p_i | i = 1, \ldots, K\}$, it can be shown that the $2K \times 1$ vector of the estimated sample quantiles for a bivariate distribution, $\hat{Q}_{TZ}(P) = [\hat{Q}_{TX}(P), \hat{Q}_{TY}(P)]'$ has an asymptotic normal distribution as given in Lemma 1.

Lemma 1 *Under Assumptions 1 and 2, as $T \to \infty$,*

$$\sqrt{T}[\hat{Q}_{TZ}(P) - Q_Z(P)] \xrightarrow{d} N(0, \Lambda), \tag{4.7}$$

where

$$\Lambda = D^{-1}\mathbf{v}(D')^{-1},$$

$$D = diag[f_X(Q_X(p_1)), \cdots, f_X(Q_X(p_K)), f_Y(Q_Y(p_1)), \cdots f_Y(Q_Y(p_K))]',$$

and

$$\mathbf{v} = \lim_{T \to \infty} E(\mathbf{mm}'),$$

where $\mathbf{m} = \frac{1}{T}[\mathcal{F}_X \mathcal{F}_Y]'$ *with* $\mathcal{F}_j = [(F_{Tj}(Q_j(p_1)) - p_1), \cdots, (F_{Tj}(Q_j(p_K)) - p_K)]$, $j = X, Y$.

Proof. See Sen (1972).□

When T and P are suppressed for simplicity, $Q_{TX}(P) = Q_X$, $Q_{TY}(P) = Q_Y$, and $Q_{TZ}(P) = Q_Z$, Given the simplified notation, the above lemma simply provides the following result:

$$\sqrt{T}(\hat{Q}_Z - Q_Z) \xrightarrow{d} N(0, \Lambda), \quad as\ T \longrightarrow \infty.$$

Lemma 1 is a useful limiting distribution for two sets of quantile points which are estimated for two associated time-dependent stochastic processes. The elements of Λ are fairly complicated and depend on the serial correlation structure of the data. While it is possible to construct consistent estimates of Λ using kernel density estimates and Newey-West style truncation arguments, it is also possible (and much less cumbersome) to employ a bootstrap resampling algorithm to estimate the elements of Λ. In this context, it is important to note that the resampling algorithm must replicate the dependence structure in the data for the bootstrap estimates to be consistent. For stationary ϕ-mixing random variables, the MBB estimation developed by Kunsch (1989) and Liu and Singh (1992) will provide consistent estimates of Λ.

As the test-statistics presented below are actually based on differences in sample quantiles, it is sufficient to estimate the variance-covariance matrix of $\hat{Q}_X - \hat{Q}_Y$. Since $Q_X - Q_Y$ can be expressed as HQ_Z, where H is a matrix such that $HQ_Z = Q_X - Q_Y$, the variance-covariance matrix of $\hat{Q}_X - \hat{Q}_Y$ is given by $\frac{1}{T}H\Lambda H'$. The various estimators of $\frac{1}{T}H\Lambda H'$ are given later.

4.2.2 Test for First-Degree Stochastic Dominance

To test XD_1Y, i.e., to test H_0: $Q_X - Q_Y \geq 0$ against H_a: $Q_X - Q_Y \ngeq 0$, it is useful to employ a version of the general Wald test for equality and inequality restrictions developed by Kodde and Palm (1986) and Wolak (1989a and 1989b). The following lemma provides a framework for both FSD and SSD tests, and, is a special case of the Kodde and Palm result.

Lemma 2 *If Q_Z can be consistently estimated by $\hat{Q}_Z$ based on a sample of size T such that $\sqrt{T}(\hat{Q}_Z - Q_Z) \xrightarrow{d} N(0,\Lambda)$, then for testing H_0: $h(Q_Z) \geq 0$ against H_a: $h(Q_Z) \ngeq 0$, the test-statistic D is defined as $D = \parallel \hat{\gamma} - \tilde{\gamma} \parallel_\Sigma = (\hat{\gamma} - \tilde{\gamma})'\Sigma^{-1}(\hat{\gamma} - \tilde{\gamma})$, where $\hat{\gamma} = \sqrt{T}h(\hat{Q}_Z)$ and $\tilde{\gamma} = \sqrt{T}h(\tilde{Q}_Z)$. $\hat{\gamma}$ is an unrestricted estimator and has large sample variance-covariance matrix $\Sigma = (\partial h/\partial Q'_Z)\Lambda(\partial h'/\partial Q_Z)$. $\tilde{\gamma}$ is a restricted estimator solving $\min_\gamma(\hat{\gamma} - \gamma)'\Sigma^{-1}(\hat{\gamma} - \gamma)$ subject to the constraint $\gamma \geq 0$. D has a large sample distribution*

$$sup_{\gamma\geq 0}Pr(D \geq q|\Sigma) = \sum_{i=0}^{K} Pr[\chi^2(K-i) \geq q]W(K,i,\Sigma)$$

with W denoting the probability that i of the K elements of $\tilde{\gamma}$ are strictly positive, and q denoting the critical value.

Proof. As shown in Kodde and Palm, a vector of parameters of interest Q_Z is formulated in terms of K independent continuous functions $h(Q_Z)$, which are differentiable in some open neighborhood of the true parameters Q_Z. The hypothesis to be tested can be written as H_0: $h_1(Q_Z) = 0$, and $h_2(Q_Z) \geq 0$ against H_a: $h_1(Q_Z) \neq 0$, and $h_2(Q_Z) \ngeq 0$. The dimensions of the partition of $h(Q_Z)$ into $h_1(Q_Z)$ and $h_2(Q_Z)$ are r and $K - r$ respectively.

It is assumed that Q_Z can be consistently estimated by $\hat{Q}_Z$ such that $\sqrt{T}(\hat{Q}_Z - Q_Z) \xrightarrow{d} N(0, \Lambda)$. Thus, by the delta method, $\sqrt{T}(h(\hat{Q}_Z) - h(Q_Z)) \xrightarrow{d} N(0, \Sigma)$, where $\Sigma = (\partial h / \partial Q'_Z)\Lambda(\partial h' / \partial Q_Z)$.

The functions of parameters, $h(Q_Z)$, can be transformed into new parameter vectors $\gamma = (\gamma'_1, \gamma'_2)'$ and $\hat{\gamma} = (\hat{\gamma}'_1, \hat{\gamma}'_2)'$, where $\gamma_i = \sqrt{T}h_i(Q_Z)$ and $\hat{\gamma}_i = \sqrt{T}h_i(\hat{Q}_Z)$.

It is shown that for H_0: $\gamma_1 = 0$, and $\gamma_2 \geq 0$ against $H_a : \gamma_1 \neq 0$, and $\gamma_2 \ngeq 0$, the Wald test-statistic is

$$\begin{aligned} D &= \| \hat{\gamma} - \tilde{\gamma} \|_\Sigma \\ &= \hat{\gamma}'_1 \Sigma_{11}^{-1} \hat{\gamma}_1 \\ &\quad + (\hat{\gamma}_2 - \tilde{\gamma}_2 - \Sigma_{21}\Sigma_{11}^{-1}\hat{\gamma}_1)'(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}(\hat{\gamma}_2 - \tilde{\gamma}_2 - \Sigma_{21}\Sigma_{11}^{-1}\hat{\gamma}_1). \end{aligned}$$

where $\hat{\gamma}$ is an unrestricted consistent estimator, $\tilde{\gamma}_1 = 0$ and $\tilde{\gamma}_2$ is the solution of

$$\min_{\gamma_2 \geq 0} (\hat{\gamma}_2 - \gamma_2 - \Sigma_{21}\Sigma_{11}^{-1}\hat{\gamma}_1)'(\Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12})^{-1}(\hat{\gamma}_2 - \gamma_2 - \Sigma_{21}\Sigma_{11}^{-1}\hat{\gamma}_1).$$

For the maximum under the null hypothesis the large sample distribution is

$$sup_{\gamma_2 \geq 0} Pr(D \geq q|\Sigma) = \sum_{i=0}^{K-r} Pr[\chi^2(K - i) \geq q] W(K - r, i, \Sigma_{22} - \Sigma_{21}\Sigma_{11}^{-1}\Sigma_{12}),$$

with W denoting the probability that i of the $K - r$ elements of $\tilde{\gamma}_2$ are strictly positive, and q denoting the critical value.

Now consider the hypotheses H_0: $\gamma_2 \geq 0$ and H_a: $\gamma_2 \ngeq 0$. This sets $r = 0$. The dimension of γ_2 is K instead of $K - r$. In this case, γ_1 is not in the parameter space, and Σ_{12}, Σ_{11} and Σ_{21} become irrelevant. Simply rename γ_2 as γ, $\hat{\gamma}_2$ as $\hat{\gamma}$, $\tilde{\gamma}_2$ as $\tilde{\gamma}$, and Σ_{22} as Σ. Then we have the following result.

For H_0: $h(Q_Z) \geq 0$ against H_a: $h(Q_Z) \ngeq 0$, the test-statistic D is defined as $D = \| \hat{\gamma} - \tilde{\gamma} \|_\Sigma = (\hat{\gamma} - \tilde{\gamma})'\Sigma^{-1}(\hat{\gamma} - \tilde{\gamma})$, where $\hat{\gamma} = \sqrt{T}h(\hat{Q}_Z)$ and $\tilde{\gamma} = \sqrt{T}h(\tilde{Q}_Z)$.

$\hat{\gamma}$ is an unrestricted estimator and has the large sample variance-covariance matrix $\Sigma = (\partial h/\partial Q_Z')\Lambda(\partial h'/\partial Q_Z)$. $\tilde{\gamma}$ is a restricted estimator solving $\min_\gamma(\hat{\gamma}-\gamma)'\Sigma^{-1}(\hat{\gamma}-\gamma)$ subject to the constraint $\gamma \geq 0$. D has a large sample distribution

$$sup_{\gamma\geq 0}Pr(D \geq q|\Sigma) = \sum_{i=0}^{K} Pr[\chi^2(K-i) \geq q]W(K,i,\Sigma)$$

with W denoting the probability that i of the K elements of $\tilde{\gamma}$ are strictly positive, and q denoting the critical value. When $i = K$, $Pr[\chi^2(0) \geq q] = 0$, for $q > 0$.

Wolak (1989a and 1989b) provides the same result in a different context.□

The upper and lower bounds for the critical values for testing inequality restrictions are provided by Kodde and Palm (1986). The reason for computing the upper- and lower-bonds for the critical value is that computing the weights W can be nontrivial. The computing of weights involves evaluation of K-multiple integrals, and closed forms are only available for small K.[2] Kodde and Palm (1986) provide a partial solution to this problem by computing the upper- and lower-bound critical values that do not require computation of the weights. These bounds are given by

$$\alpha_l = \frac{1}{2}Pr(\chi_1^2 \geq q_l), \tag{4.8}$$

and

$$\alpha_u = \frac{1}{2}Pr(\chi_{K-1}^2 \geq q_u) + \frac{1}{2}Pr(\chi_K^2 \geq q_u), \tag{4.9}$$

where q_l and q_u are the lower- and upper-bounds, respectively, for the critical values of the test-statistic. These are reproduced in Tables 4.1 and 4.2. A lower bound for the critical value is obtained by choosing a significance level α and setting degrees of freedom (df) equal to one. An upper bound for the critical value is obtained by choosing a significance level α and setting df equal to K. Decision rules based on the statistic D are: if D exceeds the upper bound value, reject H_0; and if D is smaller than the lower bound value, do not reject H_0. If D is in the inconclusive region,

[2]See Kudo (1963) for the exact computation.

then the weights W in the distribution can be determined numerically, and D can be compared with the critical value corresponding to the chosen significance level α. D is the basic form of the test-statistics for FSD and SSD.

Theorem 3 *Under Assumptions 1 and 2, the variance-covariance matrix of* $\hat{Q}_X - \hat{Q}_Y$ *is* $\frac{1}{T}H\Lambda H'$, *where* Λ *is given by Lemma 1. Under* $H_0 : Q_X - Q_Y \geq 0$, *the test-statistic for FSD,* c_1, *is given by:*

$$c_1 = \Delta'[\frac{1}{T}H\hat{\Lambda}H']^{-1}\Delta, \tag{4.10}$$

where $\Delta = [(\hat{Q}_X - \hat{Q}_Y) - (\tilde{Q}_X - \tilde{Q}_Y)]$; $\hat{Q}_X$, $\hat{Q}_Y$, *and* $\hat{\Lambda}$ *are the unrestricted estimates while* $\tilde{Q}_X$ *and* $\tilde{Q}_Y$ *are the restricted estimates minimizing*

$$[(\hat{Q}_X - \hat{Q}_Y) - (Q_X - Q_Y)]'[\frac{1}{T}H\hat{\Lambda}H']^{-1}[(\hat{Q}_X - \hat{Q}_Y) - (Q_X - Q_Y)] \tag{4.11}$$

$$s.t. \;\; (Q_X - Q_Y) \geq 0.$$

The test-statistic c_1 *is asymptotically distributed as a weighted sum of* χ^2 *random variables with different degrees of freedom i.e.*

$$\begin{aligned} & sup_{(Q_X - Q_Y) \geq 0} \quad Pr[c_1 \geq q | \frac{1}{T}H\Lambda H'] \\ = & \sum_{i=0}^{K} Pr[\chi^2(K-i) \geq q] W[K, i, \frac{1}{T}H\Lambda H']. \end{aligned} \tag{4.12}$$

The decision rules based on the statistic c_1 *are the same as those for the statistic* D *in Lemma 2.*

Proof: This result is a consequence of Lemmas 1 and 2. □

The test-statistic, c_1, employs the unrestricted estimates $\hat{Q}_X - \hat{Q}_Y$ and $\frac{1}{T}H\hat{\Lambda}H'$, and the restricted estimates $\tilde{Q}_X - \tilde{Q}_Y$ which are estimated by solving the restricted nonlinear optimization problem in (4.11).

4.2.3 Test for Second-Degree Stochastic Dominance

The SSD null hypothesis may be written as: H_0: $\Psi_X - \Psi_Y \geq 0$ against H_a: $\Psi_X - \Psi_Y \not\geq 0$.

Definition 4 *The cumulative quantile generator, B, is defined as a $(K \times K)$ lower triangular matrix with every non-zero element equal to unity, i.e.,*

$$B = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 1 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & 1 & 1 & \cdots & 1 & 0 \\ 1 & 1 & 1 & \cdots & 1 & 1 \end{bmatrix}.$$

Given K, B premultiplies $\hat{Q}_j$ $(j = X, Y)$ yielding a K-variate vector of cumulative sample quantiles, i.e.,

$$\begin{aligned} B\hat{Q}_j &= [\sum_{i=1}^{p_1} \hat{Q}_j(p_i), \ldots, \sum_{i=1}^{p_K} \hat{Q}_j(p_i))' \\ &= (\hat{\Psi}_j(p_1), \ldots, \hat{\Psi}_j(p_K))' \\ &= \hat{\Psi}_j, \end{aligned} \tag{4.13}$$

where $p_0 = 0$, $p_K = 1$ and $p_{i+1} - p_i = p_{j+1} - p_j$ for all $i, j = 0, 1, \ldots, K-1$.

Theorem 4 *Under Assumptions 1 and 2, the variance-covariance matrix of $B(\hat{Q}_X - \hat{Q}_Y)$ is $\frac{1}{T}BH\Lambda H'B'$, where Λ is given by Lemma 1. Under H_0: $B(Q_X - Q_Y) \geq 0$, the test-statistic for SSD, c_2, is given by:*

$$c_2 = (B\Delta)'[\frac{1}{T}BH\hat{\Lambda}H'B']^{-1}(B\Delta), \tag{4.14}$$

where $B\Delta = B[(\hat{Q}_X - \hat{Q}_Y) - (\tilde{Q}_X - \tilde{Q}_Y)]$. $\hat{Q}_X$, $\hat{Q}_Y$, and $\hat{\Lambda}$ are the unrestricted estimators while $\tilde{Q}_X$ and $\tilde{Q}_Y$ are the restricted estimators minimizing

$$\{B[(\hat{Q}_X - \hat{Q}_Y) - (Q_X - Q_Y)]\}'\{\frac{1}{T}BH\hat{\Lambda}H'B'\}^{-1}\{B[(\hat{Q}_X - \hat{Q}_Y) - (Q_X - Q_Y)]\} \tag{4.15}$$

$$s.t. \;\; B(Q_X - Q_Y) \geq 0.$$

c_2 is asymptotically distributed as the weighted sum of χ^2 random variables i.e.:

$$sup_{[B(Q_X - Q_Y)] \geq 0} \quad Pr[c_2 \geq q | \frac{1}{T} BH\Lambda H'B']$$
$$= \sum_{i=0}^{K} Pr[\chi^2(K-i) \geq q] W[K, i, \frac{1}{T} BH\Lambda H'B']. \tag{4.16}$$

The decision rules based on the statistic c_2 are the same as those for the statistic D in Lemma 2.

Proof: This result is derived from Lemmas 1 and 2, Definition 4 acting as a linear transformation. □

As with the FSD test-statistic, c_2 employs the unrestricted estimates $B(\hat{Q}_X - \hat{Q}_Y)$ and $\frac{1}{T} BH\hat{\Lambda}H'B'$, and the restricted $B(\tilde{Q}_X - \tilde{Q}_Y)$, which are estimated by solving the restricted nonlinear optimization problem in (4.15). It should be noted that Δ in (4.10) and Δ in (4.14) are not identical because the two are computed from different restricted optimization procedures. A simple example will serve to clarify this point. Let Δ be a 2×1 vector, which can be written as:

$$(\hat{Q}_X - \hat{Q}_Y) - (\tilde{Q}_X - \tilde{Q}_Y) = \begin{bmatrix} \hat{d}_1 - \tilde{d}_1 \\ \hat{d}_2 - \tilde{d}_2 \end{bmatrix}.$$

In (4.10), $\begin{bmatrix} \tilde{d}_1 \\ \tilde{d}_2 \end{bmatrix}$ is chosen by minimizing one objective function subject to the constraint $\begin{bmatrix} \tilde{d}_1 \\ \tilde{d}_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$, i.e.,

$$\tilde{d}_1 \geq 0, \tag{4.17}$$

and

$$\tilde{d}_2 \geq 0. \tag{4.18}$$

However, in (4.14), $\begin{bmatrix} \tilde{d}_1 \\ \tilde{d}_2 \end{bmatrix}$ is chosen by minimizing another objective function subject to the constraint $\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} \tilde{d}_1 \\ \tilde{d}_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}$. If $B = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$, the constraint becomes

$$\begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} \tilde{d}_1 \\ \tilde{d}_2 \end{bmatrix} \geq \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \text{ i.e.,}$$

$$\tilde{d}_1 \geq 0, \tag{4.19}$$

and

$$\tilde{d}_1 + \tilde{d}_2 \geq 0. \tag{4.20}$$

Clearly, restrictions (4.17) and (4.18) are different from restrictions (4.19) and (4.20). The former restricts $\tilde{d}_1$ and $\tilde{d}_2$ to be nonnegative, while the latter restricts $\tilde{d}_1$ to be nonnegative but not $\tilde{d}_2$.

In empirical work, a choice of K must be made and generally this should be an integer greater than or equal to 10. If K is too small, the comparison will be made based on relatively large intervals in the support. This may blur the true relation between the two risky prospects. If K is chosen to be a value greater than 10, say 20, the comparison between the quantiles (in the test for FSD) or the cumulative quantiles (in the test for SSD) is clearly based on smaller intervals in the support than those corresponding to the choice of $K = 10$. It must be noted that the choice of K will affect the two risky prospects symmetrically in terms of intervals on which the comparison is made. This will not change the fundamental relation between the two compared risky prospects. On the other hand, the upper and lower bounds computed by Kodde and Palm (1986) are given for df from 1 to 40 (see Tables 4.1 and 4.2). If df is chosen to be beyond this range, the table has to be extended. Finally, it should be noted that since the sample quantiles from a given individual sample will be positively correlated, there will be some point at which increasing K will produce no gain in power.

4.2.4 Moving-Block Bootstrap (MBB) Estimation

In order to construct FSD and SSD test-statistics, consistent estimates of $\frac{1}{T}H\Lambda H'$ are required. Given the complexity of Λ, a resampling procedure that is computationally convenient and can provides consistent estimates should be used.

The MBB is one of such procedures. The MBB estimator of the variance-covariance matrix for $(\hat{Q}_X - \hat{Q}_Y)$, $\frac{1}{T}H\Lambda H'$, must be used for both c and c_2 . The standard bootstrap randomly resamples with replacement from a given sample of observations, as many times, N, as necessary. The variance of the statistic of interest is then estimated using the sample variance of the statistics calculated over the bootstrap replications. Rather than resampling individual observations, the moving block bootstrap resamples blocks of observations with replacement from the original data set. Let $\{Z_t\}_{t=1}^{T}$ be a finite sample of a sequence of stationary ϕ -mixing random vectors. Let $\hat{Q}_X - \hat{Q}_Y$ be the estimator of the true population parameter of interest, $Q_X - Q_Y$. Denote the moving blocks as $B_1, \cdots, B_{T-b+1}$, where b is the size of each block and B_j stands for the block consisting of b consecutive observations starting from Z_j, i.e., $B_j = \{Z_j, Z_{j+1}, \ldots, Z_{j+b-1}\}$. For each moving block bootstrap replication, a $\hat{Q}_{X_s} - \hat{Q}_{Y_s}$ can be computed. If the resampling takes place N times, an empirical sampling distribution of $\hat{Q}_X - \hat{Q}_Y$ can be constructed, along with various statistics associated with the distribution. In particular, the variance-covariance matrix of $\hat{Q}_X - \hat{Q}_Y$ can be computed using the sample variance of $\hat{Q}_{X_s} - \hat{Q}_{Y_s}$, $s = 1, 2, \ldots, N$.

Consistency for the MBB estimation is achieved if the number of observations in each block, b, approaches infinity with T in such a way that the number of moving blocks, $k = [T/b]$, also approaches infinity with T. In general, larger values for b are necessary to capture stronger dependence. Another practical concern is the choice for the number of bootstrap replications, N. According to Efron (1982), and Efron and Tibshirani (1986), it is quite adequate for N to be in the range from 50 to 200 for estimating of a variance.

4.3 Tests under I.I.D. and Association Conditions

4.3.1 Assumptions

If the weak dependence assumption is replaced with the *i.i.d.* assumption, two samples may still be statistically dependent, and hence be characterized by a joint bivariate distribution. This simplification is suitable for certain income distribution data where the same individuals are sampled at different points in time.

In order to develop the tests, Assumption 1 is replaced with Assumption 3.

Assumption 3 *The sample observations from the joint distribution, $F_{XY}(x,y)$, of X and Y are i.i.d..*

Under Assumptions 2 and 3, the joint asymptotic distribution of quantiles from a bivariate distribution is given by Lemma 3.

Lemma 3 *Under Assumptions 2 and 3, as $T \to \infty$,*

$$\sqrt{T}(\hat{Q}_Z - Q_Z) \xrightarrow{d} N(0, \Lambda), \tag{4.21}$$

where the elements of Λ are given by

$$\frac{F_{js}(Q_j(p_i)), Q_s(p_t)) - p_i p_t}{f_j(Q_j(p_i)) f_s(Q_s(p_t))}$$

for $j, s = X, Y \quad (j \neq s)$, and $i, t = 1, 2, \ldots, K$;

$$\frac{p_i(1 - p_t)}{f_j(Q_j(p_i)) f_j(Q_j(p_t))}$$

for $j = X, Y \ (j = s)$, and $i, t = 1, 2, \ldots, K$.

Proof. See Siddiqui (1960) and Weiss (1964).□

The elements of $\widehat{\Lambda}$ can be consistently estimated using a variety of nonparametric density estimations. [see Appendix for nonparametric estimation].

4.3.2 Tests for First-Degree and Second-Degree Stochastic Dominance

The test-statistic for first-degree stochastic dominance under Assumptions 2 and 3 is given in Corollary 1.

Corollary 1 *Under Assumptions 2 and 3 , the variance-covariance matrix of* $\hat{Q}_X - \hat{Q}_Y$ *is* $\frac{1}{T}H\Lambda H'$, *where* Λ *is given by Lemma 3. Under* $H_0 : Q_X - Q_Y \geq 0$, *the test-statistic for FSD,* c_1^d, *is defined as*

$$c_1^d = \Delta'[\frac{1}{T}H\hat{\Lambda}H']^{-1}\Delta. \tag{4.22}$$

The decision rules based on the statistic c_1^d *are the same as those for the statistic D in Lemma 2.*

Proof. Straightforward from Theorem 3 and Lemma 3. □

The test-statistic for second-degree stochastic dominance under Assumptions 2 and 3 is given in Corollary 2.

Corollary 2 *Under Assumptions 1 and 2, the variance-covariance matrix of* $B(\hat{Q}_X - \hat{Q}_Y)$ *is* $\frac{1}{T}BH\Lambda H'B$, *where* Λ *is given by Lemma 3. Under* H_0: $B(Q_X - Q_Y) \geq 0$,*the statistic for SSD,* c_2^d, *is defined as*

$$c_2^d = (B\Delta)'[\frac{1}{T}BH\hat{\Lambda}H'B']^{-1}(B\Delta). \tag{4.23}$$

The decision rules based on the statistic c_2^d *are the same as those for the statistic D in Lemma 2.*

Proof. Straightforward from Theorem 4 and Lemma 3. □

In summary, the test-statistics c_1^d and c_2^d are used in the FSD and SSD tests, respectively. These are distributed as a weighted sum of χ^2 random variables.

4.4 Tests under I.I.D. and Independence Conditions

4.4.1 Assumptions

When the two samples are drawn independently, Lemma 3 can be further simplified. This simplification is suitable for certain income distribution data; namely, samples that are independently selected, each sample having data that are *i.i.d.*.

In order to develop the tests, Assumption 2 is replaced with Assumption 4.

Assumption 4 *The sample observations for each random variable are i.i.d. and X and Y are independent.*

Under this assumption, only the univariate distribution functions, $F_X(x)$ and $F_Y(y)$, instead of the joint distribution function, $F(x,y)$, are needed. Under Assumptions 2 and 4, the asymptotic distribution for the quantile function estimates of each random variable is given in Lemma 4.

Lemma 4 *Under Assumptions 2 and 4, as $T \to \infty$,*

$$\sqrt{T}(\hat{Q}_Z - Q_Z) \xrightarrow{d} N(0, \Lambda), \tag{4.24}$$

where $\Lambda = \begin{bmatrix} \Lambda_X & 0 \\ 0 & \Lambda_Y \end{bmatrix}$ *with elements in* Λ_j, $j = X, Y$ *given by*

$$\frac{p_i(1-p_t)}{f_j(Q_j(p_i))f_j(Q_j(p_t))}$$

for $j = X, Y$, *and* $i, t = 1, 2, \ldots, K$.

Proof. See Mosteller (1946), Rao (1965), Wilks (1962), and Beach and Davidson (1983). □

It should be noted that $\frac{1}{T}H\Lambda H' = \frac{1}{T}(\Lambda_X + \Lambda_Y)$ where Λ is as shown in Lemma 4.

4.4.2 Tests for First-Degree and Second-Degree Stochastic Dominance

The test-statistic for first-degree stochastic dominance under Assumptions 2, and 4 is given in Corollary 3.

Corollary 3 *Under Assumptions 2, and 4, the variance-covariance matrix of* $(\widehat{Q}_X - \widehat{Q}_Y)$ *is* $\frac{1}{T}H\Lambda H' = \frac{1}{T}(\Lambda_X + \Lambda_Y)$, *where* Λ *is given in Lemma 4. Under* H_0: $Q_X - Q_Y \geq 0$, *the test-statistic for FSD,* c_1^i, *is defined as*

$$c_1^i = \Delta'[\frac{1}{T}H\widehat{\Lambda}H']^{-1}\Delta. \tag{4.25}$$

The decision rules based on the statistic c_1^i *are the same as those for the statistic D in Lemma 2.*

Proof. Straightforward from Theorem 3 and Lemma 4. □

The test-statistic for second-degree stochastic dominance under Assumptions 2, and 4 is given in Corollary 4.

Corollary 4 *Under Assumptions 2, and 4, the variance-covariance matrix of* $B(\widehat{Q}_X - \widehat{Q}_Y)$ *is* $\frac{1}{T}BH\Lambda H'B' = \frac{1}{T}B(\Lambda_X + \Lambda_Y)B'$, *where* Λ *is given in Lemma 4. Under* H_0: $B(Q_X - Q_Y) \geq 0$, *the test-statistic for SSD,* c_2^i, *is defined as*

$$c_2^i = (B\Delta)'[\frac{1}{T}BH\widehat{\Lambda}H'B']^{-1}(B\Delta). \tag{4.26}$$

The decision rules based on the statistic c_2^i *are the same as those for the statistic D in Lemma 2.*

Proof. Straightforward from Theorem 4 and Lemma 4. □

In summary, the test-statistics c_1^i and c_2^i are used in the FSD and SSD tests, respectively, under Assumptions 2, and 4. These are distributed as a weighted sum of χ^2 random variables. It should be noted that c_1^d and c_2^d are general forms of c_1^i and c_2^i, respectively. Further, c_1 and c_2 are more general forms of c_1^d and c_2^d, respectively.

4.5 Concluding Remarks

It has been noted that the existing test procedures for stochastic dominance suffer from many restrictions. Two prominent difficulties of those procedures are that the dominance relation is not properly specified under the null hypothesis, and that the test procedures cannot accommodate complex data structures. This chapter proposes new distribution-free tests under very general conditions. The test procedure can test the dominance relation properly and has desirable asymptotic distribution. The test-statistics under Assumptions 1 and 2 are c_1 and c_2, as shown in Theorems 3 and 4. When the condition of weak dependence are replaced with the *i.i.d.* condition, and association still remains, two other simplified tests are proposed. The test-statistics under Assumptions 2 and 3 are c_1^d and c_2^d, as demonstrated in Corollaries 1 and 2. When the condition of association is replaced with the independence conditions, two even simpler tests are derived. The test-statistics under Assumptions 2 and 4 are c_1^i and c_2^i, as given in Corollaries 3 and 4.

4.6 Appendix: Nonparametric Density-Quantile Estimation

Consistent estimates of the variance-covariance matrices of sample quantiles in Corollaries 1, 2, 3, and 4, can be provided by various nonparametric procedures. The description given below is based on the method of kernels. Useful descriptions are also provided by Silverman (1986), Ullah (1988), and Izenman (1992).

Let $\{x_t\}_{t=1}^T$ be a sample of observations drawn from a continuous univariate distribution with probability density function f. Let $\hat{f}$ be the kernel estimator of f with kernel K_ζ and window width ζ, i.e.,

$$\hat{f}(w) = \frac{1}{T\zeta}\sum_{t=1}^{T} K_\zeta\left(\frac{w - x_t}{\zeta}\right), \tag{4.27}$$

in which w is a reference point and $w \in [a,b]$. [3] K_ζ is a continuous and symmetric function which satisfies $\int_a^b K_\zeta(w)dw = 1$ and $K_\zeta \geq 0$. While various forms of K_ζ are possible, the Gaussian K_ζ is used here because much empirical evidence suggest that the choice of kernel is not crucial to the estimates obtained. The univariate Gaussian kernel is given by

$$K_\zeta(w - x_t) = \frac{1}{(2\pi)^{.5}} e^{-.5\left(\frac{w-x_t}{\zeta}\right)^2}. \tag{4.28}$$

More crucial is the choice of the window width ζ, since this determines the amount of smoothing undertaken and may affect the consistency of the estimator.

The nonparametric density estimator has desirable asymptotic properties. [4] Using the L_2 approach, Parzen (1962) shows that, under regularity conditions on K_ζ, the univariate kernel estimator is both asymptotically unbiased and asymptotically normal. Using the L_1 approach, Devroye (1983) proves that, under regularity condi-

[3]It is much more general if $w \in (-\infty, +\infty)$. We keep the notations consistent throughout. Thus, $w \in [a, b]$ is defined.

[4]Finite-sample properties of nonparametric density are available for special situations [see Fryer (1976)].

tions on K_ζ, the kernel estimator $\hat{f}$ is a strongly consistent estimator of f, without any conditions on f.[5]

It is also possible to integrate the density estimates to get distribution estimates, and to get quantile estimates with proper transformation. [6]

The estimated or empirical density, distribution, and quantile functions can be used to explore the relations among risky prospects. Furthermore, these estimates can be used in the new test-statistics for FSD and SSD.

The window width can be (i) selected manually with the individual researcher's judgement; (ii) estimated by minimizing the mean integrated square error; or (iii) determined by the cross-validations. The methods for selecting ζ are given below.

When the ζ is selected manually, the choices are based on subjective judgement. The window width ζ can also be chosen by minimizing the mean integrated square error,

$$MISE(\hat{f}) = \int E(\hat{f}(w) - f(w))^2 dw. \tag{4.29}$$

When a Gaussian kernel is being used, then the window width obtained from the above minimization would be

$$\zeta_{opt} = \left(\frac{4}{3}\right)^{\frac{1}{5}} \sigma T^{-\frac{1}{5}} = 1.06\sigma T^{-\frac{1}{5}}. \tag{4.30}$$

In computing, σ can be replaced with $\hat{\sigma}$. Alternatively, using a robust measure of spread, $\zeta_{opt} = 0.79RT^{-\frac{1}{5}}$, where R is the interquartile range of the data. The window width ζ can also be estimated by either least squares cross-validation ($LSCV$) or maximum-likelihood cross-validation ($MLCV$). $LSCV$ is to select a ζ such that the

[5]See also Devroye and Gyorfi (1985), Chapter 8.

[6]The asymptotic properties of the density estimator will be maintained for the estimators of the distribution function as well the quantile function, due to the Mann and Wald (1943) theorem [See also Rao (1965), p.104]. In practical terms, the empirical distribution estimator, $\hat{F}(w) = \int_a^w \hat{f}(t)dt$, is a consistent estimator of $F(w)$ [Reiss (1981) and Yang (1985) provide some discussion of the kernel density estimator of the distribution function]. The empirical quantile estimator, $\hat{Q}(p) = \hat{F}^{-1}(p)$, is also a consistent estimator of $Q(p)$.

function $LSCV(\zeta)$ is minimized. $LSCV(\zeta)$ is defined as

$$LSCV(\zeta) = \int \hat{f}^2 - 2T^{-1}\sum_t \hat{f}_{-t}(x_t), \tag{4.31}$$

where the *leave-one-out* density estimate $\hat{f}_{-t}(x_t)$ is constructed from all the data points except x_t. $\hat{f}_{-t}(x_t)$ is defined as

$$\hat{f}_{-t}(x_t) = (T-1)^{-1}\zeta^{-1}\sum_{s\neq t} I_{ts}K_\zeta(x_t - x_s), \tag{4.32}$$

where I_{ts} is an indicator that is one if observations t and s fall in the same category, zero otherwise. $MLCV$ is to select a ζ such that the function $MLCV(\zeta)$ is maximized. $MLCV(\zeta)$ is defined as

$$MLCV(\zeta) = T^{-1}\sum_{t=1}^{T} \log \hat{f}_{-t}(x_t). \tag{4.33}$$

Table 4.1: Upper and Lower Bounds for the Critical Value for Jointly Testing Equality and Inequality Restrictions

d.f.	$\alpha = .25$	$\alpha = .10$	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$	$\alpha = .001$
1	0.455	1.642	2.706	3.841	5.412	6.635	9.500
2	2.090	3.808	5.138	6.483	8.273	9.643	12.810
3	3.475	5.528	7.045	8.524	10.501	11.971	15.357
4	4.776	7.097	8.761	10.384	12.583	14.045	17.612
5	6.031	8.574	10.371	12.103	14.325	15.968	19.696
6	7.257	9.998	11.911	13.742	16.074	17.791	21.666
7	8.461	11.383	13.401	15.321	17.755	19.540	23.551
8	9.648	12.737	14.853	16.856	19.348	21.232	25.370
9	10.823	14.067	16.274	18.345	20.972	22.879	27.133
10	11.987	15.377	17.670	19.824	22.525	24.488	28.856
11	13.142	16.670	19.045	21.268	24.049	26.056	30.542
12	14.289	17.949	20.410	22.691	25.549	27.616	32.196
13	15.430	19.216	21.742	24.096	27.026	29.143	33.823
14	16.566	20.472	23.096	25.484	28.485	30.649	35.425
15	17.696	21.718	24.384	26.856	29.927	32.136	37.005
16	18.824	22.956	25.689	28.219	31.353	33.607	38.566
17	19.943	24.186	26.983	29.569	32.766	35.063	40.109
18	21.060	25.409	28.268	30.908	34.167	36.505	41.636
19	22.174	26.625	29.545	32.237	35.556	37.935	43.148
20	23.285	27.835	30.814	33.557	36.935	39.353	44.646

Source: Kodde and Palm (1986).

Table 4.2: Upper and Lower Bounds for the Critical Value for Jointly Testing Equality and Inequality Restrictions (Continued)

$d.f.$	$\alpha = .25$	$\alpha = .10$	$\alpha = .05$	$\alpha = .025$	$\alpha = .01$	$\alpha = .005$	$\alpha = .001$
21	24.394	29.040	32.077	34.869	38.304	40.761	46.133
22	25.499	30.240	33.333	36.137	39.664	42.185	47.607
23	26.602	31.436	34.583	37.470	41.016	43.547	49.071
24	27.703	32.627	35.827	38.761	42.360	44.927	50.524
25	28.801	33.813	37.066	40.045	43.696	46.299	51.986
26	29.898	34.996	38.301	41.324	45.026	47.663	53.403
27	30.992	36.176	39.531	42.597	46.349	49.020	54.830
28	32.085	37.352	40.756	43.365	47.667	50.371	56.248
29	33.176	38.524	41.977	45.128	48.978	51.715	57.660
30	34.266	39.694	43.194	46.387	50.286	53.054	59.064
31	35.354	40.861	44.408	47.641	51.585	54.386	60.461
32	36.440	42.025	45.618	48.891	52.881	55.713	61.852
33	37.525	43.186	46.825	50.137	54.172	57.035	63.237
34	38.609	44.345	48.029	51.379	55.459	58.352	64.616
35	39.691	45.501	49.229	52.618	56.742	59.665	65.989
36	40.773	46.655	50.427	53.853	58.020	60.973	67.357
37	41.853	47.808	51.622	55.085	59.295	62.276	68.720
38	42.932	48.957	52.814	56.313	60.566	63.576	70.078
39	44.010	50.105	54.003	57.539	61.833	64.871	71.432
40	45.087	51.251	55.190	58.762	63.097	66.163	72.780

Source: Kodde and Palm (1986).

Chapter 5

Finite Sample Properties of the Proposed Tests

5.1 Introduction

All of the test-statistics described in Chapter 4 are based on asymptotic theory. It is also important to investigate whether the asymptotic distributions of these test-statistics represent a good approximations to the true finite sample distributions for sample sizes that are typical from an economic perspective. This chapter investigates the finite sample properties of the test-statistics for first- and second-degree stochastic dominance using Monte Carlo simulations. The data are generated from parametric distribution functions such as the normal and lognormal distributions. The former represent symmetric types of distribution functions which are used as close approximations for financial asset returns; and the latter represent the asymmetric type of distribution function which may be used as a good approximation of income distributions. The dominance relation between two data generation processes (DGP's) are known in advance by specifying the parameters of the DGP's. The sampling distributions of the test-statistics can be obtained through Monte Carlo simulations. These distributions provide the information on the performance of the test-statistics under the specified conditions.

The Monte Carlo simulations are implemented for the test-statistics c_1^i, c_2^i, c_1^d, and c_2^d. While c_1^i and c_2^i are suitable for *i.i.d.* samples from independent random variables, c_1^d, and c_2^d are appropriate for *i.i.d.* samples from random variables that are dependent. The test-statistics, c_1^i and c_2^i, are given by Corollaries 1 and 2; and the test-statistics, c_1^d, and c_2^d, are illustrated in Corollaries 3 and 4. The test-statistics are distributed as a weighted sum of χ^2 distributions under the null hypotheses. The tests are similar in construct but have different variance-covariance matrices depending on the data structures.

The size and power of the test-statistics are good indicators of their performance. But the test-statistics developed herein have very complex asymptotic distributions and hence have no statistical tables readily available. Generally, the empirical size and power can be computed for a test-statistic for a single pair of parameters from its sampling distribution if one theoretical critical value is chosen for an appropriate significance level, say 5%.[1] It should be noted that the test-statistics developed in Chapter 4 involve the multiple comparision of K pairs of parameters and a composite null of a special form thereby making problematic the computation of unique critical values. For this reason, upper and lower bounds for critical values are used. Thus empirical power must be calculated for the upper (rejection) bound and empirical size must be calculated for the lower (acceptance) bound. If the precise critical value, that is in the interval bounded by the lower and upper bounds, were provided, the precise empirical size might be smaller and the precise empirical power might be greater. Because the complexity involved in solving the problem, this issue is left for future research. In this setting, the Monte Carlo simulations are designed to evaluate:

1. the performance of the test-statistics for various DGP's; and

[1] Alternatively, to evaluate the test-statistic for a finite sample, the sampling distribution of the test-statistic can also be integrated to get an empirical critical value for a certain significance level. Then this empirical critical value can be compared with the theoretical critical value for the same significance level.

2. the effects of misspecification on the test-statistics, in particular, the impact on a test-statistic when it is applied to mismatched data.

5.2 Data Generating Processes and Experiment Design

To investigate the finite sample performance of the test-statistics, Monte Carlo simulations with following characteristics are used:

1. a pair of samples of size T are randomly selected from specific parametric distribution(s);

2. the dominance test-statistic (for either FSD or SSD) is then computed using nonlinear optimization with a set of constraints imposed on the parameters;

3. this process is repeated N times and the test-statistics are computed each time, thereby yielding a sampling distribution; and

4. the basic statistics such as the mean, standard deviation, minimum, and maximum, and the empirical size and power of the test-statistics are computed from the sampling distribution.

Several factors must be selected for each simulation:

1. Normal and Lognormal Distributions
   The normal and lognormal distribution functions are used. Clearly, to evaluate the finite sample properties of the proposed test-statistics, the DGP's must be specified for the random variables X and Y. These distribution functions are chosen for relevance and feasibility. Relevance requires that the parametric distributions can represent the key features of distributions tested. Normal distributions are used to approximate financial asset returns; lognormal distributions are are probably the simplest models for income distribution. Feasibility

demands that the parameters of the parametric distribution should reveal the true dominance relation between two random variables. Both normal and lognormal distributions can be characterized by their mean and variance, and hence are suitable candidate distributions.

2. Parameters—means, variances, and covariances

   The parameters of the parametric functions are chosen such that a dominance relation between two DGP's is well specified. A normally distributed random variable, say X, with mean μ_X and variance σ_X^2 is denoted by $X \sim N(\mu_X, \sigma_X^2)$. A random sample can be generated using the normal random variable generator with specific μ_X and σ_X^2. Let random variable, say X', be a specific transformation of X, like $X' = e^X$. X' is lognormally distributed with mean $e^{\mu_X + \sigma_X^2/2}$ and variance $e^{2\mu_X + \sigma_X^2}(e^{\sigma_X^2} - 1)$, i.e., $X' = e^X \sim LN[e^{\mu_X + \sigma_X^2/2}, e^{2\mu_X + \sigma_X^2}(e^{\sigma_X^2} - 1)]$. To generate a random sample from the lognormally distributed random variable X', a random sample from the normally distributed random variable X is first generated, and then transformed into a random sample of the corresponding lognormally distributed random variable X'. It is important to note that μ_X and σ_X^2 affect both the mean and variance of X' and care should be taken in selecting μ_X and σ_X^2. When X and Y are associated and assumed to be distributed jointly, the covariance, σ_{XY}, must also be specified in addition to the mean and variance of each marginal distribution.

   To apply the test for FSD, two cases are considered: When $\mu_X \geq \mu_Y$ and $\sigma_X = \sigma_Y$, X dominates Y in the first-degree. When $\mu_X < \mu_Y$ and $\sigma_X = \sigma_Y$, X does not dominate Y in the first-degree. To apply the test for SSD, again two cases are considered: When $\mu_X \geq \mu_Y$ and $\sigma_X \leq \sigma_Y$, X dominates Y in the second-degree. When $\mu_X < \mu_Y$ and $\sigma_X > \sigma_Y$, X does not dominate Y in the second-degree.

3. Sample size (T)

   The sample size is directly related the issue of finite sample properties of test-statistics. For this purpose, Monte Carlo simulations for samples of size 400 are conducted, because this size approaches the sample size in the empirical application in Chapter 5. Of course, 400 is much smaller than the usual sample size that income distribution data have.

4. Number of point estimates on a quantile function (K)
   The issue of optimal spacing of order statistics has been the subject of much research [see, for example, David (1981)]. Given the optimal spacing of order statistics, the central tendency and dispersion can be uniquely and efficiently determined. From this literature, spacing into twenty order statistics at most is characteristically recommended. For the simulations, the number of point estimates is therefore chosen to be 20.

5. Number of replications (N)
   In general, the choice of N is *ad hoc* depending on the criteria of the experiment. Davidson and MacKinnon (1993) suggest that the number of replications in the Monte Carlo simulation be 1000, 2000, 5000, and 10000. They also say that it may be as small as 50 if estimation is very time-consuming and accurate results are not needed. N in our simulation is chosen to be 1000. The simulation reported here is extremely time-consuming because each replication requires standard iterative nonlinear estimation. In general, 1000 replications take an IBM compatible 486 computer (33MHz and 50 MHz) more than 24 hours to complete.

6. Size of the test (α)
   The size of the test is denoted α. To evaluate the empirical size of the test, the empirical distribution of the test-statistic should be generated. From this empirical distribution, the empirical size of the test can be computed using the theoretical critical value that is related to α. In this paper, α is set equal to 0.05. If the number of the point estimates on a quantile function is 20, at

5% significance level, the lower- and upper-bounds of the critical value for the test-statistics are 2.706 and 30.814, respectively.

Two types of simulations are implemented. The first type of simulations is for c_1^i and c_2^i with the DGP's which have *i.i.d.* samples from independent random variables X and Y. The second type of simulations is for c_1^d and c_2^d with the DGP's which have *i.i.d* samples from dependent random variables X and Y. The size and power of the tests are evaluated empirically for various DGP's. In addition, the experiments are conducted for the misspecification of c_1^d and c_2^d as c_1^i and c_2^i, respectively.

5.3 Simulation Evidence

In the following, the simulation results for c_1^i and c_2^i are reported first. Among thirty-six experiments, the first 18 experiments for FSD and SSD based on the normal are reported in Tables 5.1 and 5.2. The remaining 18 experiments for FSD and SSD based on the lognormal distributions are reported later.

In Table 5.1, the Monte Carlo simulation results for FSD are reported. The first five DGP's (see the first five rows) for X and Y satisfy the FSD relation. Cases 6–9 do not satisfy the inequality condition specified under the null hypothesis. Thus, cases 1–9 represent a spectrum of cases from FSD to non-FSD. The mean of the test-statistic c_1^i changes from 0 to 735.001. The sampling distributions of these test-statistics are of our primary concern. When the DGP's follow the FSD relation, such as cases 1–4, the empirical probability of accepting the null hypothesis ranges from 1.00 to 0.982 if the critical value at the 5 % significance level is chosen based on the asymptotic distribution. In case five where the two DGP's are identical, sampling errors blur the relation so that the empirical probability of accepting the null hypothesis is 0.817 at the 5 % significance level. In Cases 6–9, there is no FSD. If the lower-bound of critical values is used, there is no evidence that the null hypothesis of FSD should be accepted in cases 7–9. The empirical power for cases 7–9 ranges from .94 to 1 except for case 6 where the two DGP's are too close for the test to have any power. For each case, the sampling distribution of the test-statistic is graphed. These are presented in Figures 5.1–5.9.

Table 5.2 presents the results of the Monte Carlo simulation for SSD. The first five DGP's (see the first five rows) for X and Y satisfy the FSD relation. Cases 6–9 do not satisfy the inequality condition specified under the null hypothesis. Thus, cases 1–9 represent a spectrum of cases from SSD to non-SSD. The mean of the test-statistic c_2^i changes from 0 to 234.676. When the DGP's have the SSD relation as in cases 1–4, the empirical probability of accepting the null hypothesis is always equal to 1.00 if the critical value at the 5 % significance level is used. In case five where the two DGP's are identical, sampling errors blur the relation so that the empirical probability of accepting the null hypothesis is 0.914 at the 5 % significance level. Cases 6–9 are for the DGP's that do not satisfy the condition of SSD. When the lower-bound of critical values is used, there is no evidence for the null hypothesis of SSD to be accepted in cases 6–9. The empirical power for cases 7–9 are ranging from 0.699 to 1 except for case 6 where the two DGP's are too close to allow the test has any power. For each case, the sampling distribution of the test-statistic is graphed. These are presented in Figures 5.10–5.18.

The remaining eighteen experiments for FSD and SSD are based on the log-normal distributions. For these experiments, eight lognormally distributed random variables are generated. The random variables used are numbered with bold type numbers from **1** to **8**, and their parametric specifications are provided in Table 5.3. The first column of Table 5.3 lists the identification numbers of the 8 random variables. The second and third columns give the specifications of mean and variance of the corresponding normally distributed random variables. The dominance relations among the lognormally distributed random variables are obvious, i.e., the variables with smaller identification numbers dominate ones with greater identification numbers.

Among eighteen experiments conducted for the lognormally distributed random variables, 1–8, the first nine experiments are for tests for FSD and the other nine experiments are used to evaluate the tests for SSD. A summary of the experimental results are reported in Tables 5.4 and 5.5.

In the upper part of Table 5.4, the first five cases where X and Y satisfy the FSD relation are listed in the first five rows. Cases 6–9 (see rows 6–9 in the upper part of the table) deviate from the FSD relation by construction. In the lower part of Table 5.4, the mean, standard error, minimum, and maximum of the test-statistics, c_1^i, are given in columns 4–7, respectively. The percentages of the test-statistics that are less than 2.706 and greater than 30.814 are listed in columns 8 and 9, respectively. The mean of the test-statistics, c_1^i, changes from 2.222 in row one to 210.314 in row nine. When the DGP's have the FSD relation, such as in cases 1–4, the empirical probability for the test-statistics less than 2.706 range from 0.828 to 0.871. In case 5 where two DGP's are identical, sampling errors lead to 66.4% of the test-statistics being less than the lower-bound 2.706. Cases 6–9 are the cases where there is no FSD. If the lower bound is used, there is little probability that the test-statistics are less than the lower-bound 2.706.

In Table 5.5, the first five cases where X and Y satisfy the SSD relation are given in the first five rows in the upper part of the table. Cases 6–9 (see rows 6–9 in the upper part of the table) deviate from the SSD relation by construction. In the lower part of Table 5.5, the mean, standard error, minimum, and maximum of the test-statistics, c_2^i, are given in columns 4–7, respectively. The percentages of the test-statistics that are less than 2.706 and greater than 30.814 are listed in columns 8 and 9, respectively. The mean of the test-statistics, c_2^i, changes from 0.329 in row one to 205.408 in row nine. When the DGP's allow the SSD, such as in cases 1–4, the empirical probability for the test-statistics less than 2.706 ranges from 0.979 to 0.963. In case 5 where two DGP's are identical, sampling errors lead to 81.4% of the test-statistics being less than the lower-bound 2.706. Cases 6–9 are the cases where there is no SSD. If the lower bound is used, there is a low probability that the test-statistics are less than the lower-bound 2.706.

To see the size and power of the tests clearly, let the empirical and theoretical size be α^e and α, respectively. While α is selected, α^e and β^e are computed from the sampling distributions of the test-statistics. Note that the asymptotic distribution of the test-statistics is a weighted sum of χ^2 distributions and hence is not easy to compute. However, the lower- and upper-bounds of the critical value, q_l and q_u, are tabulated for a given α, as given in the previous chapter. If the test-statistics are less than q_l, then the null hypothesis of a dominance relation should not be rejected at the corresponding significance level. If the test-statistics are greater than q_u, then the null hypothesis of a dominance relation should be rejected at the corresponding significance level. Thus, the empirical size and power should be computed using q_l and q_u, respectively. The lower (higher) the value of α^e, the more (less) reliable the test will be. The higher (lower) the value of β^e, the greater (smaller) power the test will have.

Table 5.6 reports α^e and β^e for Table 5.1 that summarizes the FSD experiment results for various normal distributions. It is clear that the empirical size α^e's shown in column 4 are smaller than the theoretical size α in cases 1–4. Case 5 has α^e that is greater than α because the two DGP's are identical. While the power of the test, β^e, for cases 7–9, are very high, the test has no power in case 6 because the two DGP's do not differ sufficiently enough. Table 5.7 illustrates α^e and β^e for Table 5.2 that summarizes the SSD experiment results for various normal distributions. In Table 5.7, the empirical size α^e's are smaller than the theoretical size α in cases 1-4. In case 5, α^e is only slightly greater than α when the two DGP's are identical. While the power of the test, β^e, for cases 8–9, are very high, the test has lower power in case 5, and no power in case 6 because two DGP's being compared do not differ sufficiently enough. Similar patterns can be observed from Tables 5.8 and 5.9 from the FSD and SSD experiments of lognormal distributions. In general, the empirical size of the FSD or SSD test for lognormal distributions is slightly higher than that for normal distributions. When the two DGP's being compared are very close, the power of the FSD or SSD test for lognormal distributions is slightly lower than that for normal distributions.

The above are the summaries of the simulations for the test-statistics c_1^i and c_2^i using both normal and lognormal distributions. In addition, the simulation results for c_1^d and c_2^d are given in comparison to the simulation results of misspecification of c_1^d and c_2^d as c_1^i and c_2^i, respectively. The comparison is made using the bivariate normal distributions. For the DGP's, data are generated under the assumption of nonzero covariance (assume $\sigma_{XY} = 0.7$). One set of simulations takes the nonzero covariance into consideration ($\sigma_{XY} \neq 0$). Another set of simulations imposes a zero covariance restriction ($\sigma_{XY} = 0$) on estimation and inference, this represents a misspecification. The simulation results for FSD and SSD are reported in Tables 5.10, and 5.11, respectively. The top parts of the tables list the simulation results for the correctly specified test-statistics (c_1^d and c_1^d) which take the nonzero σ_{XY} into consideration; the lower parts of the tables show the simulation results for the misspecified test-statistics (c_1^i and c_1^i) which impose the zero restriction on σ_{XY}. As expected for the same DGP's, the average values of the misspecified test-statistics are generally lower than those of the correctly specified test-statistics for both FSD and SSD. This can be explained with an example. When random variables X and Y can be characterized by a bivariate joint distribution and have a positive covariance, $\sigma_{XY} > 0$, a correctly specified test-statistic for the difference between the two means, say $\mu_X - \mu_Y = 0$, should use the sample counterpart of $\sigma_X^2 - 2\sigma_{XY} + \sigma_Y^2$. One of possible misspecified test-statistics for the same relationship under identical conditions may use the sample counterpart of $\sigma_X^2 + \sigma_Y^2$, thereby imposing the restriction $\sigma_{XY} = 0$. Thus, the misspecified test of similar nature uses a greater sample variance and hence has a smaller value.

5.4 Concluding Remarks

The empirical size and power are computed for various cases in Tables 5.1, 5.2, 5.4, and 5.5. The empirical size and power are listed in Tables 5.6, 5.7, 5.8, and 5.9. In Table 5.6, cases 1–4 have empirical size smaller than the theoretical size; and cases 7–9 have empirical power close or equal to one. For marginal cases, such as cases 5 and 6, while the empirical size of the test for case 5 is higher, the test has no power for case 6. Table 5.7 has a similar pattern. When the tests are applied to the lognormal distributions (see Tables 5.8, and 5.9), a similar pattern is observed, except that the empirical size is generally greater than its theoretical counterpart for FSD. Thus, it can be concluded that the proposed tests for FSD and SSD are remarkably effective when the DGP's are significantly different. However, if the DGP's only differ marginally, the test performs reasonably well when the DGP's conform with the specification of the null hypothesis; it performs poorly when the DGP's do not conform with the specification of the null hypothesis. When the tests are not properly specified for the data structures (such as using c_1^i and c_2^i for c_1^d and c_2^d, respectively), the misspecification will cause inaccurate statistics to be computed.

Table 5.1: Simulation Results of Test for XD_1Y: Normal Distribution

	(1)	(2)	(3)
	X	Y	Relation
(1)	$N(2.0, 1.0)$	$N(0.0, 1.0)$	XD_1Y
(2)	$N(1.0, 1.0)$	$N(0.0, 1.0)$	XD_1Y
(3)	$N(0.5, 1.0)$	$N(0.0, 1.0)$	XD_1Y
(4)	$N(0.1, 1.0)$	$N(0.0, 1.0)$	XD_1Y
(5)	$N(0.0, 1.0)$	$N(0.0, 1.0)$	XD_1Y
(6)	$N(0.0, 1.0)$	$N(0.1, 1.0)$	$X\not D_1Y$
(7)	$N(0.0, 1.0)$	$N(0.5, 1.0)$	$X\not D_1Y$
(8)	$N(0.0, 1.0)$	$N(1.0, 1.0)$	$X\not D_1Y$
(9)	$N(0.0, 1.0)$	$N(2.0, 1.0)$	$X\not D_1Y$

	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	Std. Error	Minimum	Maximum	% < 2.706	% > 30.814
(1)	0.000	0.000	0.000	0.000	1.000	0.000
(2)	0.000	0.000	0.000	0.000	1.000	0.000
(3)	0.000	0.000	0.000	0.007	1.000	0.000
(4)	0.294	0.653	0.000	4.547	0.982	0.000
(5)	1.496	1.886	0.000	14.626	0.817	0.000
(6)	5.080	4.022	0.000	23.259	0.320	0.000
(7)	50.275	13.475	17.604	107.330	0.000	0.940
(8)	188.402	28.252	112.850	291.170	0.000	1.000
(9)	735.001	65.506	552.780	980.920	0.000	1.000

Table 5.2: Simulation Results of Test for XD_2Y: Normal Distribution

	(1)	(2)	(3)
	X	Y	Relation
(1)	$N(0.5, 0.5)$	$N(0.0, 2.0)$	XD_2Y
(2)	$N(0.5, 1.0)$	$N(0.0, 2.0)$	XD_2Y
(3)	$N(0.5, 1.5)$	$N(0.0, 2.0)$	XD_2Y
(4)	$N(0.5, 1.8)$	$N(0.0, 2.0)$	XD_2Y
(5)	$N(0.0, 2.0)$	$N(0.0, 2.0)$	XD_2Y
(6)	$N(0.0, 2.0)$	$N(0.5, 1.8)$	$X\not{D}_2Y$
(7)	$N(0.0, 2.0)$	$N(0.5, 1.5)$	$X\not{D}_2Y$
(8)	$N(0.0, 2.0)$	$N(0.5, 1.0)$	$X\not{D}_2Y$
(9)	$N(0.0, 2.0)$	$N(0.5, 0.5)$	$X\not{D}_2Y$

	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	Std. Dev.	Minimum	Maximum	% < 2.706	% > 30.814
(1)	0.000	0.000	0.000	0.000	1.000	0.000
(2)	0.000	0.000	0.000	0.001	1.000	0.000
(3)	0.000	0.000	0.000	0.000	1.000	0.000
(4)	0.000	0.000	0.000	0.010	1.000	0.000
(5)	0.[illegible]25	1.332	0.000	11.697	0.914	0.000
(6)	18.288	8.018	2.608	58.603	0.000	0.074
(7)	37.585	11.706	10.574	87.340	0.000	0.699
(8)	108.786	19.885	57.655	193.760	0.000	1.000
(9)	234.676	34.506	135.890	364.310	0.000	1.000

Table 5.3: Definitions of Ten Lognormally Distributed Random Variables

Random Variable No.	μ	σ^2	$e^{\mu+\sigma^2/2}$	$e^{2\mu+\sigma^2}(e^{\sigma^2}-1)$
1	0.900	0.200	2.718	1.636
2	0.800	0.300	2.586	2.339
3	0.700	0.400	2.460	2.975
4	0.600	0.500	2.340	3.551
5	0.500	0.600	2.226	4.072
6	0.400	0.700	2.117	4.543
7	0.300	0.800	2.014	4.970
8	0.200	0.900	1.916	5.356

Table 5.4: Simulation Results of Test for XD_1Y: Lognormal Distribution

	(1)	(2)	(3)
	X	Y	Relation
(1)	1	8	$1D_1\ 8$
(2)	2	7	$2D_1\ 7$
(3)	3	6	$3D_1\ 6$
(4)	4	5	$4D_1\ 5$
(5)	5	5	$5D_1\ 5$
(6)	5	4	$5\not{D}_1\ 4$
(7)	6	3	$6\not{D}_1\ 3$
(8)	7	2	$7\not{D}_1\ 2$
(9)	8	1	$8\not{D}_1\ 1$

	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	Std. Dev.	Minimum	Maximum	% $<$ 2.706	% $>$ 30.814
(1)	2.226	6.296	0.000	120.160	0.828	0.007
(2)	2.320	9.811	0.000	148.460	0.894	0.016
(3)	1.833	9.035	0.000	174.450	0.908	0.014
(4)	1.887	6.647	0.000	98.088	0.871	0.015
(5)	3.697	10.743	0.000	217.420	0.664	0.014
(6)	8.005	6.823	0.309	75.607	0.107	0.012
(7)	35.438	9.604	5.198	86.127	0.000	0.673
(8)	95.191	17.811	40.833	195.130	0.000	1.000
(9)	210.314	32.619	124.250	325.260	0.000	1.000

Table 5.5: Simulation Results of Test for XD_2Y: Lognormal Distribution

	(1)	(2)	(3)
	X	Y	Relation
(1)	1	8	$1D_2$ 8
(2)	2	7	$2D_2$ 7
(3)	3	6	$3D_2$ 6
(4)	4	5	$4D_2$ 5
(5)	5	5	$5D_2$ 5
(6)	6	4	$5\not{D}_2$ 4
(7)	7	3	$6\not{D}_2$ 3
(8)	8	2	$7\not{D}_2$ 2
(9)	9	1	$8\not{D}_2$ 1

	(4)	(5)	(6)	(7)	(8)	(9)
	Mean	Std. Dev.	Minimum	Maximum	% < 2.706	% > 30.814
(1)	0.329	2.682	0.000	35.916	0.979	0.002
(2)	0.340	2.697	0.000	36.578	0.978	0.002
(3)	0.496	3.906	0.000	63.990	0.976	0.004
(4)	0.570	3.273	0.000	55.516	0.963	0.003
(5)	2.008	5.221	0.000	79.269	0.814	0.008
(6)	6.272	8.229	0.007	172.440	0.217	0.009
(7)	32.397	9.636	9.355	90.429	0.000	0.529
(8)	90.949	16.464	46.029	159.770	0.000	1.000
(9)	205.408	32.375	119.140	307.600	0.000	1.000

Table 5.6: Empirical Size and Power of Test for XD_1Y: Normal Distribution

	(1)	(2)	(3)	(4)	(5)
	X	Y	Relation	α^e	β^e
(1)	$N(2.0, 1.0)$	$N(0.0, 1.0)$	XD_1Y	0.000	—
(2)	$N(1.0, 1.0)$	$N(0.0, 1.0)$	XD_1Y	0.000	—
(3)	$N(0.5, 1.0)$	$N(0.0, 1.0)$	XD_1Y	0.000	—
(4)	$N(0.1, 1.0)$	$N(0.0, 1.0)$	XD_1Y	0.018	—
(5)	$N(0.0, 1.0)$	$N(0.0, 1.0)$	XD_1Y	0.183	—
(6)	$N(0.0, 1.0)$	$N(0.1, 1.0)$	$X\not{D}_1Y$	—	0.000
(7)	$N(0.0, 1.0)$	$N(0.5, 1.0)$	$X\not{D}_1Y$	—	0.940
(8)	$N(0.0, 1.0)$	$N(1.0, 1.0)$	$X\not{D}_1Y$	—	1.000
(9)	$N(0.0, 1.0)$	$N(2.0, 1.0)$	$X\not{D}_1Y$	—	1.000

Table 5.7: Empirical Size and Power of Test for XD_2Y: Normal Distribution

	(1)	(2)	(3)	(4)	(5)
	X	Y	Relation	α^e	β^e
(1)	$N(0.5, 0.5)$	$N(0.0, 2.0)$	XD_2Y	0.000	—
(2)	$N(0.5, 1.0)$	$N(0.0, 2.0)$	XD_2Y	0.000	—
(3)	$N(0.5, 1.5)$	$N(0.0, 2.0)$	XD_2Y	0.000	—
(4)	$N(0.5, 1.8)$	$N(0.0, 2.0)$	XD_2Y	0.000	—
(5)	$N(0.0, 2.0)$	$N(0.0, 2.0)$	XD_2Y	0.086	—
(6)	$N(0.0, 2.0)$	$N(0.5, 1.8)$	$X\not D_2Y$	—	0.074
(7)	$N(0.0, 2.0)$	$N(0.5, 1.5)$	$X\not D_2Y$	—	0.699
(8)	$N(0.0, 2.0)$	$N(0.5, 1.0)$	$X\not D_2Y$	—	1.000
(9)	$N(0.0, 2.0)$	$N(0.5, 0.5)$	$X\not D_2Y$	—	1.000

Table 5.8: Empirical Size and Power of Test for XD_1Y: Lognormal Distribution

	(1)	(2)	(3)	(4)	(5)
	X	Y	Relation	α^e	β^e
(1)	1	8	$1D_1$ 8	0.172	—
(2)	2	7	$2D_1$ 7	0.106	—
(3)	3	6	$3D_1$ 6	0.092	—
(4)	4	5	$4D_1$ 5	0.129	—
(5)	5	5	$5D_1$ 5	0.336	—
(6)	5	4	$5\not{D}_1$ 4	—	0.012
(7)	6	3	$6\not{D}_1$ 3	—	0.673
(8)	7	2	$7\not{D}_1$ 2	—	1.000
(9)	8	1	$8\not{D}_1$ 1	—	1.000

Table 5.9: Empirical Size and Power of Test for XD_2Y: Lognormal Distribution

	(1)	(2)	(3)	(4)	(5)
	X	Y	Relation	α^e	β^e
(1)	1	8	$1D_2$ 8	0.021	—
(2)	2	7	$2D_2$ 7	0.022	—
(3)	3	6	$3D_2$ 6	0.024	—
(4)	4	5	$4D_2$ 5	0.037	—
(5)	5	5	$5D_2$ 5	0.186	—
(6)	5	4	$5\not{D}_2$ 4	—	0.009
(7)	6	3	$6\not{D}_2$ 3	—	0.529
(8)	7	2	$7\not{D}_2$ 2	—	1.000
(9)	8	1	$8\not{D}_2$ 1	—	1.000

Table 5.10: FSD Simulation Results for Bivariate Normal Distributions

	c_1^d for $\hat{\sigma}_{XY} \neq 0$ when $\sigma_{XY} = 0.7$			
Relation	Mean	Std. Dev.	% < 2.706	% > 30.814
$N(2.0, 1.0^2)\ D_1\ N(0.0, 1.0^2)$	0.0002	0.0067	1.0000	0.0000
$N(1.0, 1.0^2)\ D_1\ N(0.0, 1.0^2)$	0.4056	7.2975	0.9970	0.0030
$N(0.5, 1.0^2)\ D_1\ N(0.0, 1.0^2)$	1.1017	6.3916	0.9540	0.0160
$N(0.0, 1.0^2)\ \not{D}_1\ N(0.5, 1.0^2)$	121.7563	21.6127	0.0000	1.0000
$N(0.0, 1.0^2)\ \not{D}_1\ N(1.0, 1.0^2)$	478.8015	56.5744	0.0000	1.0000
$N(0.0, 1.0^2)\ \not{D}_1\ N(2.0, 1.0^2)$	1878.7408	177.6798	0.0000	1.0000

	c_1^i for $\hat{\sigma}_{XY} = 0$ when $\sigma_{XY} = 0.7$			
Relation	Mean	Std. Dev.	% < 2.706	% > 30.814
$N(2.0, 1.0^2)\ D_1\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(1.0, 1.0^2)\ D_1\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(0.5, 1.0^2)\ D_1\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(0.0, 1.0^2)\ \not{D}_1\ N(0.5, 1.0^2)$	49.6783	7.5864	0.0000	0.9970
$N(0.0, 1.0^2)\ \not{D}_1\ N(1.0, 1.0^2)$	187.3682	18.5791	0.0000	1.0000
$N(0.0, 1.0^2)\ \not{D}_1\ N(2.0, 1.0^2)$	735.6198	53.2142	0.0000	1.0000

Table 5.11: SSD Simulation Results for Bivariate Normal Distributions

	c_2^d for $\widehat{\sigma}_{XY} \neq 0$ when $\sigma_{XY} = 0.7$			
Relation	Mean	Std. Dev.	% < 2.706	% > 30.814
$N(2.0, 1.0^2)\ D_2\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(1.0, 1.0^2)\ D_2\ N(0.0, 1.0^2)$	0.3741	11.8288	0.9990	0.0010
$N(0.5, 1.0^2)\ D_2\ N(0.0, 1.0^2)$	0.0176	0.5578	0.9990	0.0000
$N(0.0, 1.0^2)\ \not{D}_2\ N(0.5, 1.0^2)$	99.1427	16.7883	0.0000	1.0000
$N(0.0, 1.0^2)\ \not{D}_2\ N(1.0, 1.0^2)$	389.9762	37.6839	0.0000	1.0000
$N(0.0, 1.0^2)\ \not{D}_2\ N(2.0, 1.0^2)$	1558.7752	114.4085	0.0000	1.0000

	c_2^i for $\widehat{\sigma}_{XY} = 0$ when $\sigma_{XY} = 0.7$			
Relation	Mean	Std. Dev.	% < 2.706	% > 30.814
$N(2.0, 1.0^2)\ D_2\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(1.0, 1.0^2)\ D_2\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(0.5, 1.0^2)\ D_2\ N(0.0, 1.0^2)$	0.0000	0.0000	1.0000	0.0000
$N(0.0, 1.0^2)\ \not{D}_2\ N(0.5, 1.0^2)$	45.4145	7.8072	0.0000	0.9820
$N(0.0, 1.0^2)\ \not{D}_2\ N(1.0, 1.0^2)$	179.8408	18.0693	0.0000	1.0000
$N(0.0, 1.0^2)\ \not{D}_2\ N(2.0, 1.0^2)$	715.2807	51.8608	0.0000	1.0000

Figure 5.1: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(2.0, 1.0)$ and $Y \sim N(0.0, 1.0)$

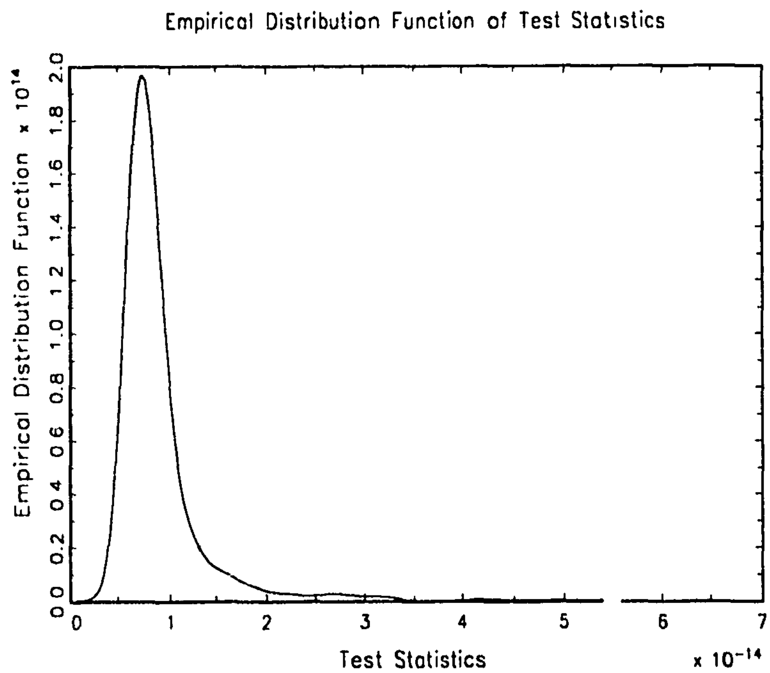


Figure 5.2: The empirical distribu\`'on of test-statistics when the null hypothesis is $XD_{1}1$, ..^{d} $X \sim N(1.0, 1.0)$ **and** $Y \sim N(0.0, 1.0)$

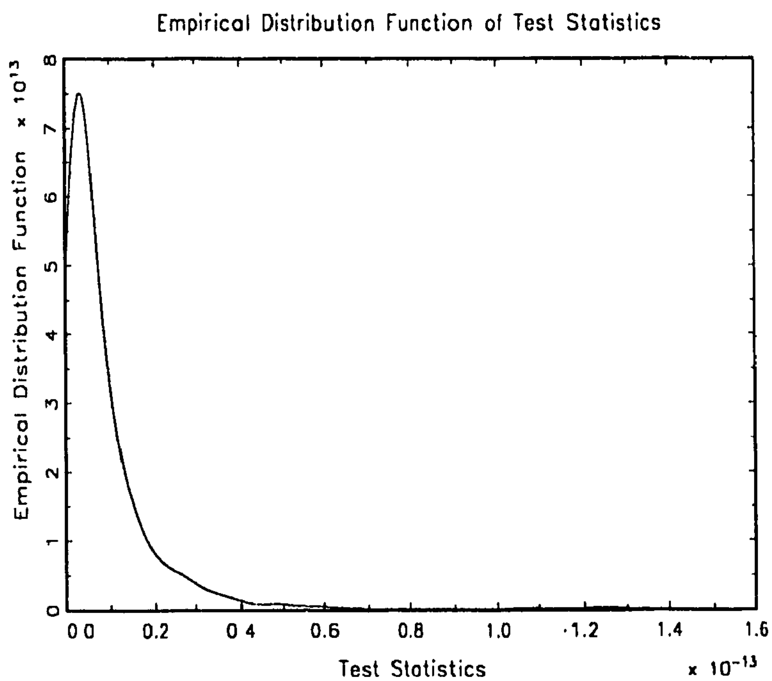


Figure 5.3: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.5, 1.0)$ and $Y \sim N(0.0, 1.0)$

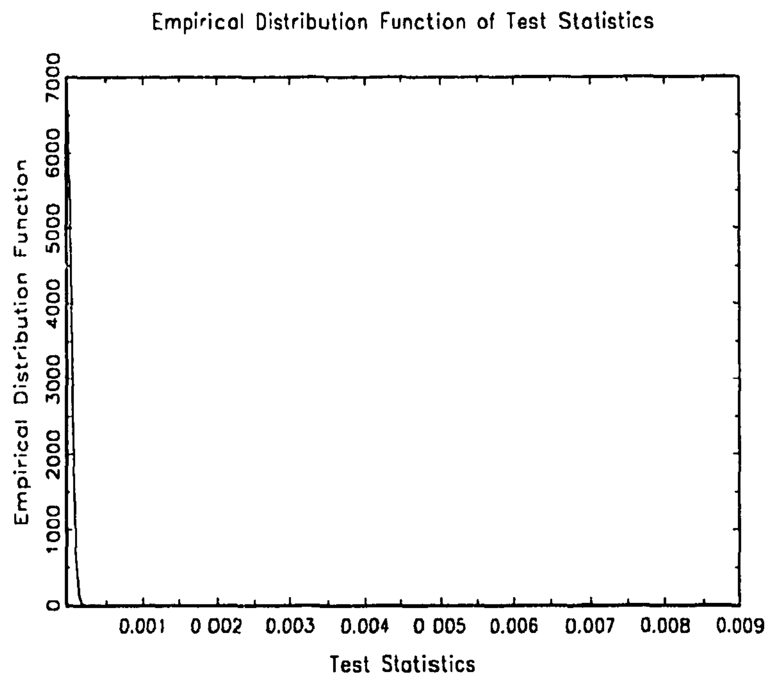


Figure 5.4: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.1, 1.0)$ and $Y \sim N(0.0, 1.0)$

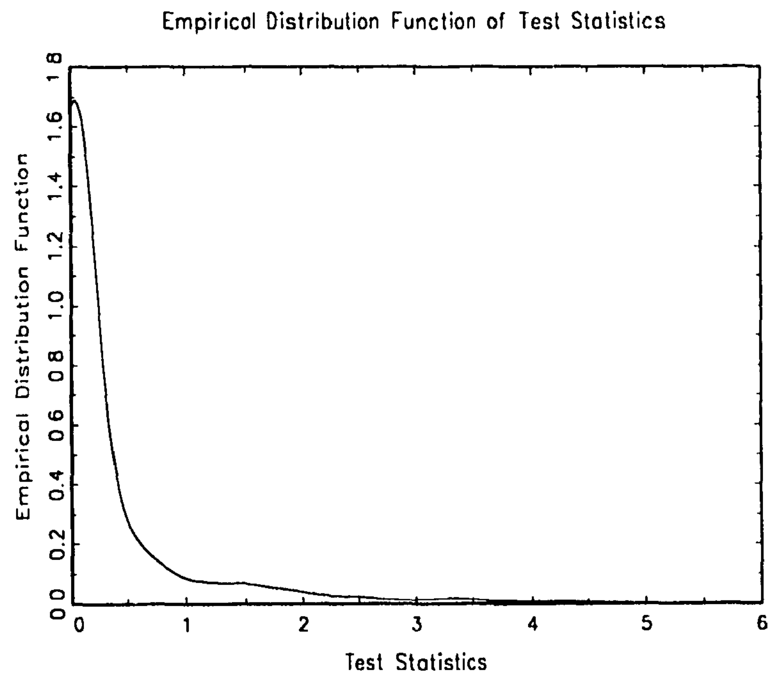


Figure 5.5: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.0, 1.0)$ and $Y \sim N(0.0, 1.0)$

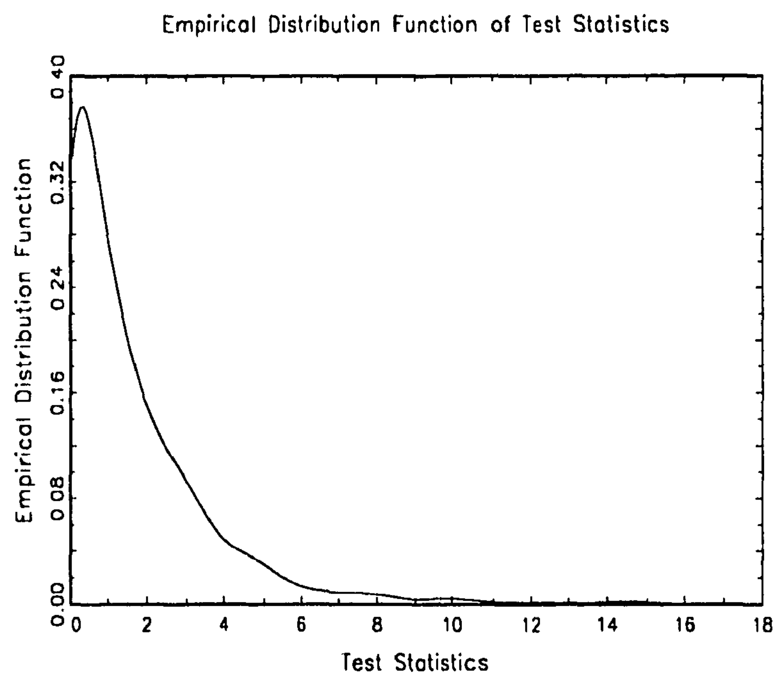


Figure 5.6: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.0, 1.0)$ and $Y \sim N(0.1, 1.0)$

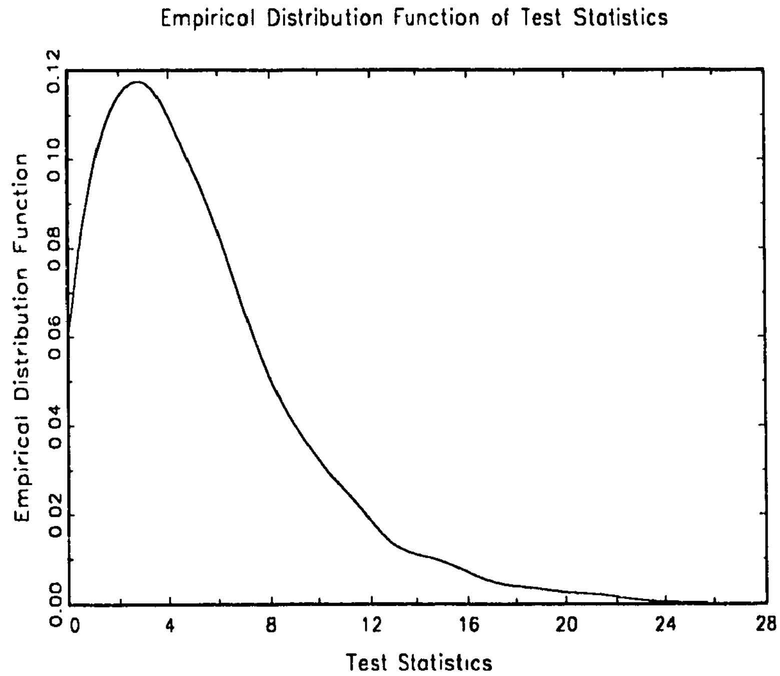


Figure 5.7: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.0, 1.0)$ and $Y \sim N(0.5, 1.0)$

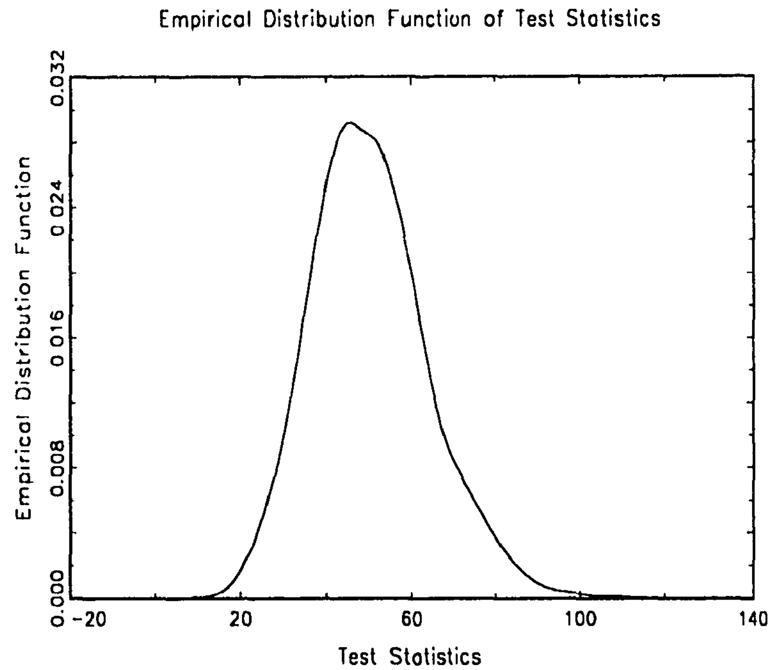


Figure 5.8: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.0, 1.0)$ and $Y \sim N(1.0, 1.0)$

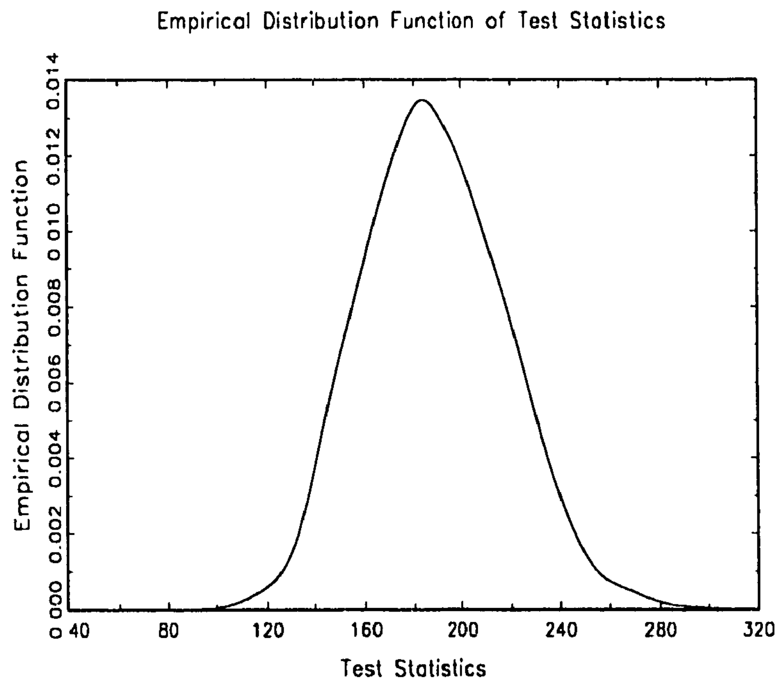


Figure 5.9: The empirical distribution of test-statistics when the null hypothesis is XD_1Y, and $X \sim N(0.0, 1.0)$ and $Y \sim N(2.0, 1.0)$

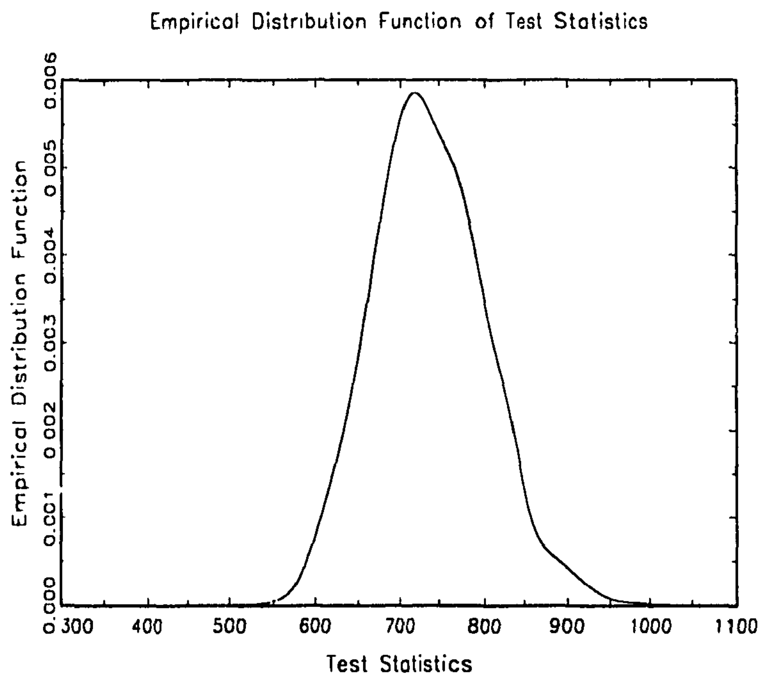


Figure 5.10: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.5, 0.5)$ and $Y \sim N(0.0, 2.0)$

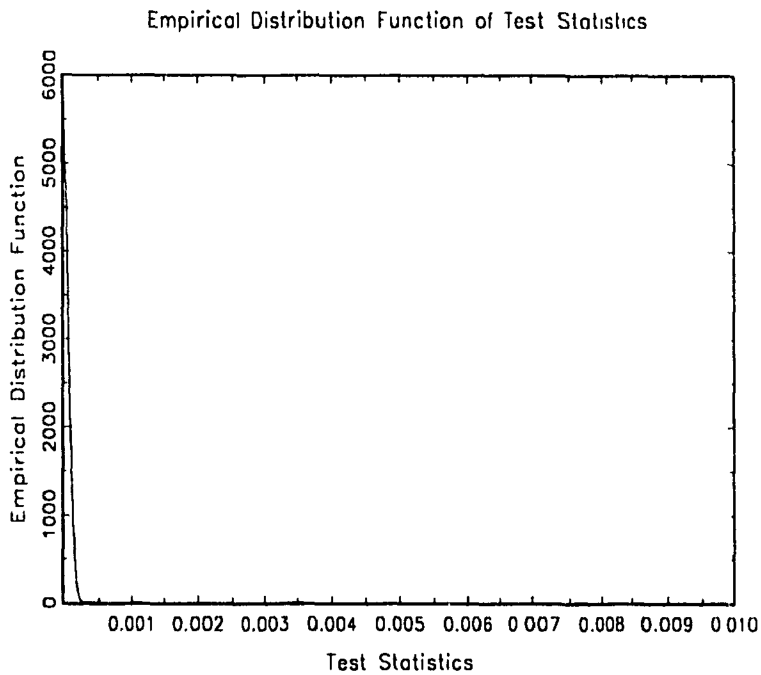


Figure 5.11: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.5, 1.0)$ and $Y \sim N(0.0, 2.0)$

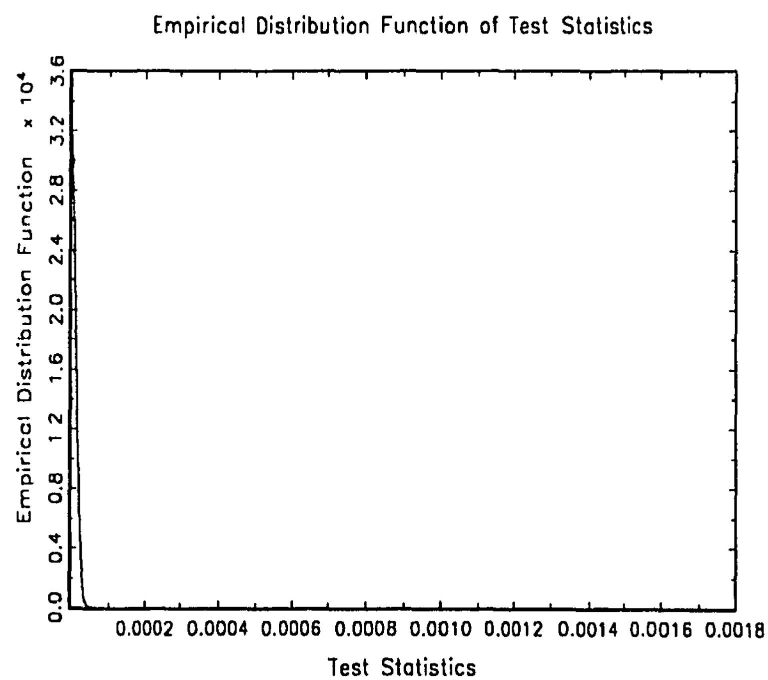


Figure 5.12: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.5, 1.5)$ and $Y \sim N(0.0, 2.0)$

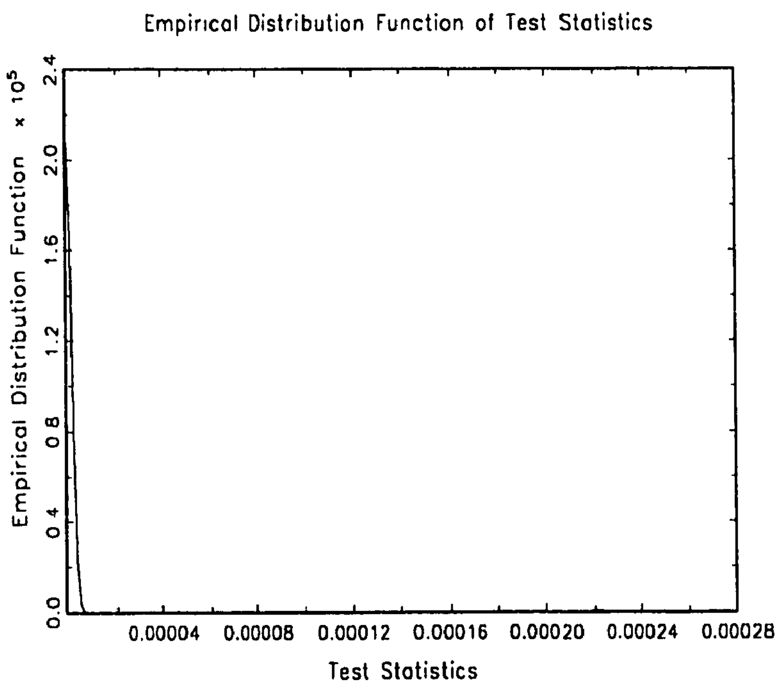


Figure 5.13: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.5, 1.8)$ and $Y \sim N(0.0, 2.0)$

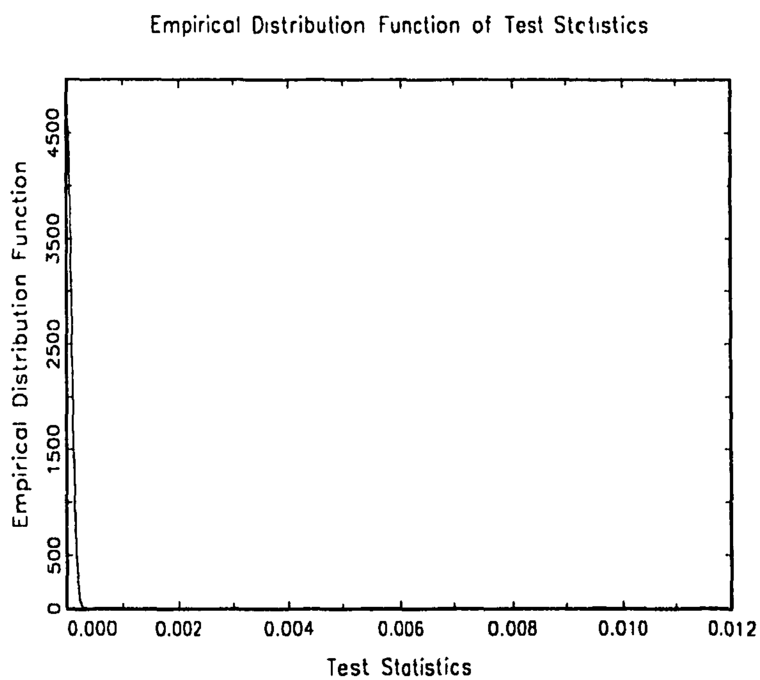


Figure 5.14: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.0, 2.0)$ and $Y \sim N(0.0, 2.0)$

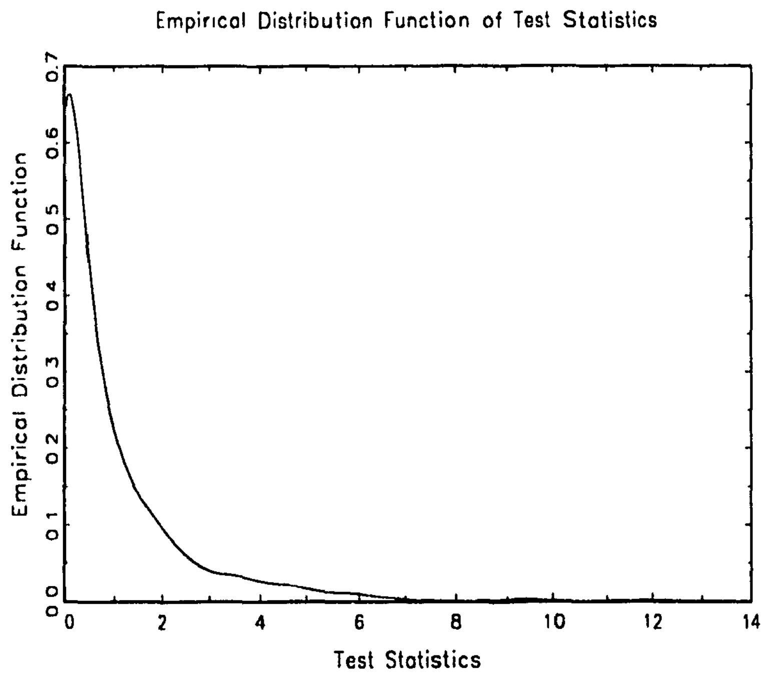


Figure 5.15: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.0, 2.0)$ and $Y \sim N(0.5, 1.8)$

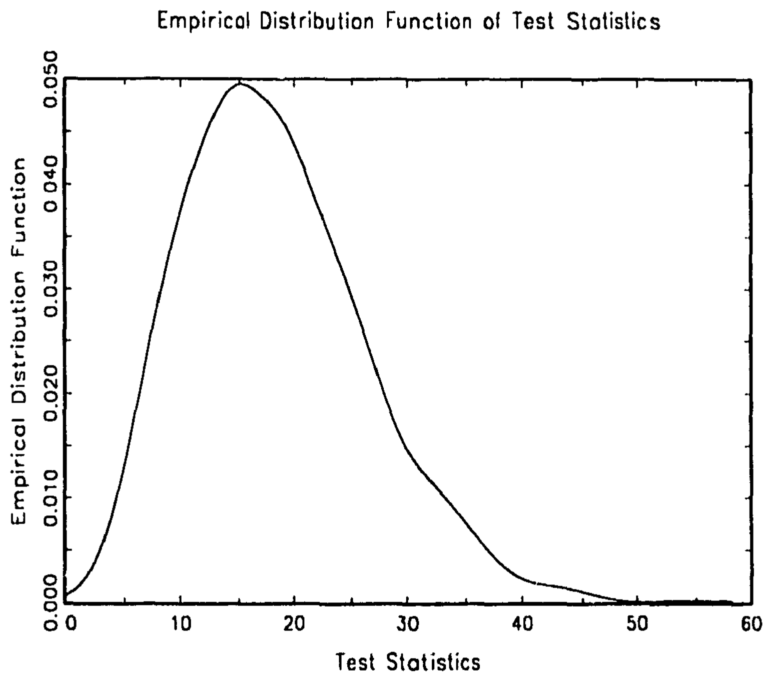


Figure 5.16: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.0, 2.0)$ and $Y \sim N(0.5, 1.5)$

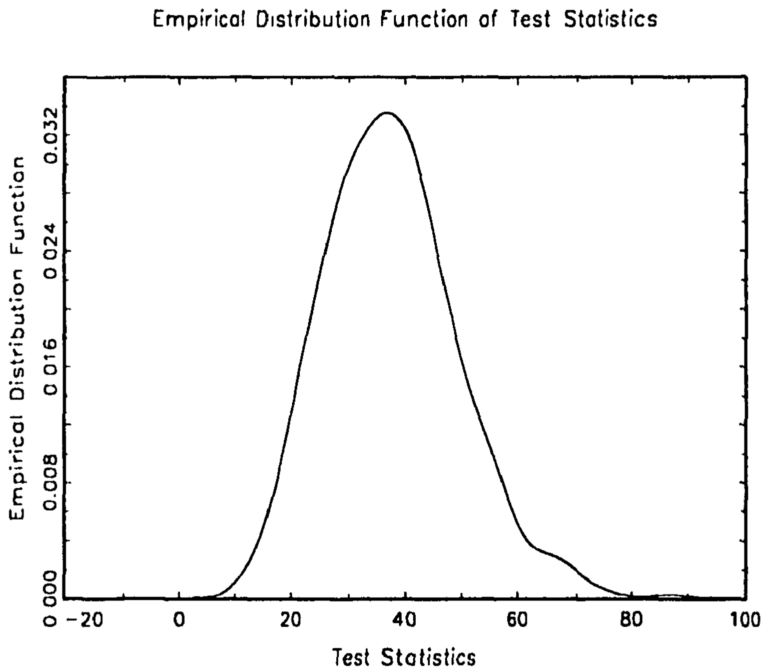


Figure 5.17: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.0, 2.0)$ and $Y \sim N(0.5, 1.0)$

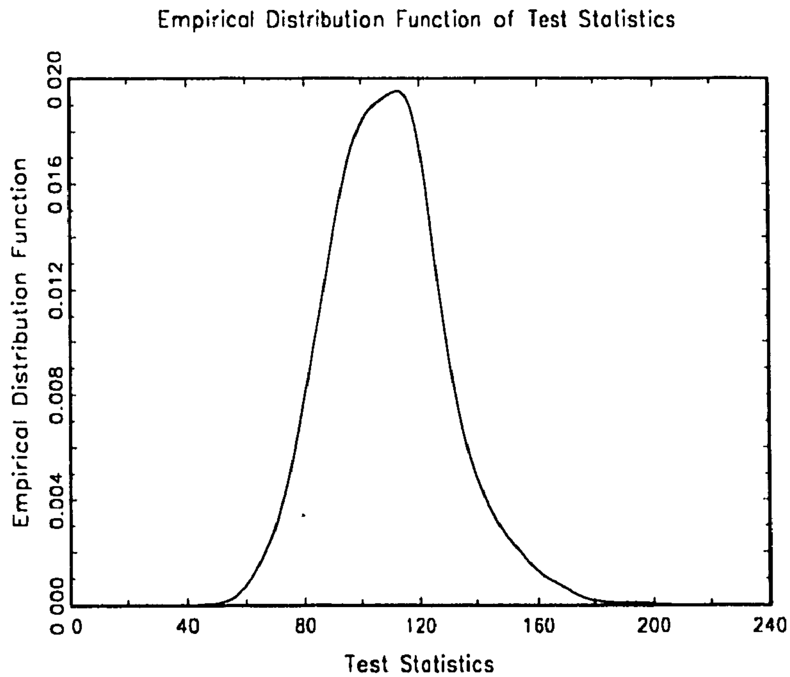


Figure 5.18: The empirical distribution of test-statistics when the null hypothesis is XD_2Y, and $X \sim N(0.0, 2.0)$ and $Y \sim N(0.5, 0.5)$

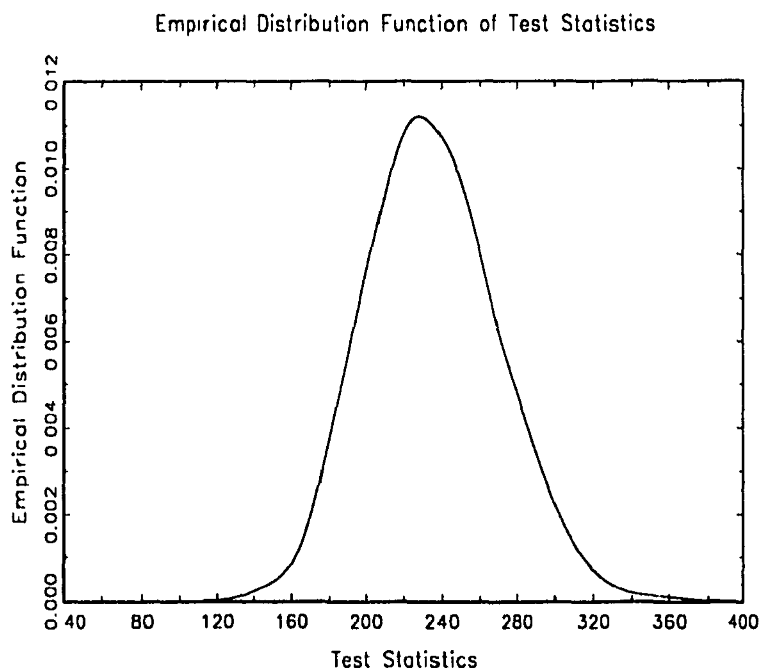


Figure 5.19: The empirical distribution of test-statistics when the null hypothesis is $\mathbf{1}D_1\mathbf{8}$

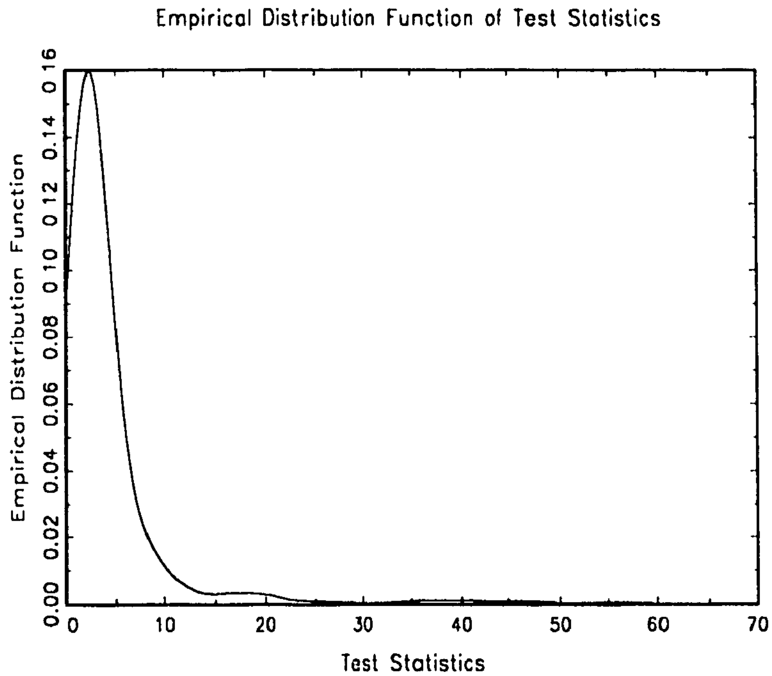


Figure 5.20: The empirical distribution of test-statistics when the null hypothesis is $2D_17$

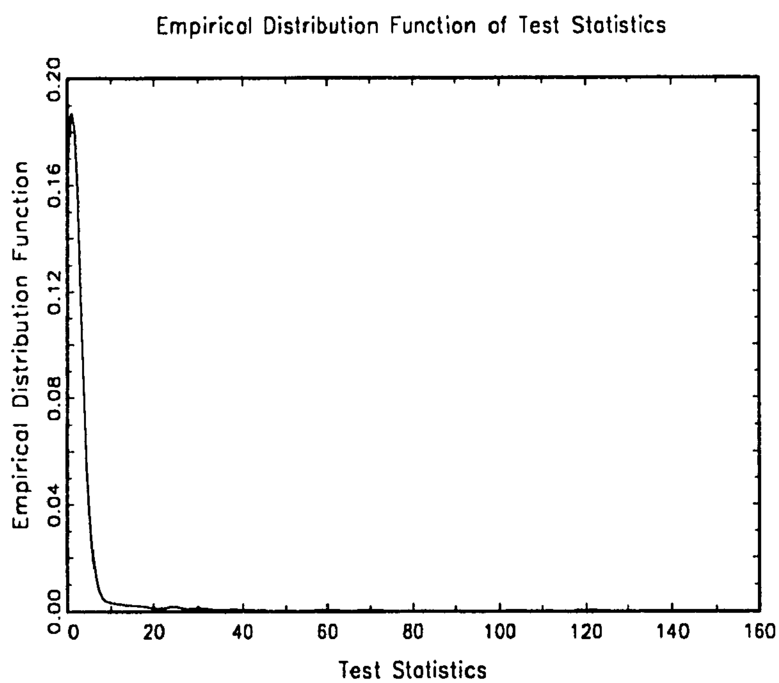


Figure 5.21: The empirical distribution of test-statistics when the null hypothesis is $3D_16$

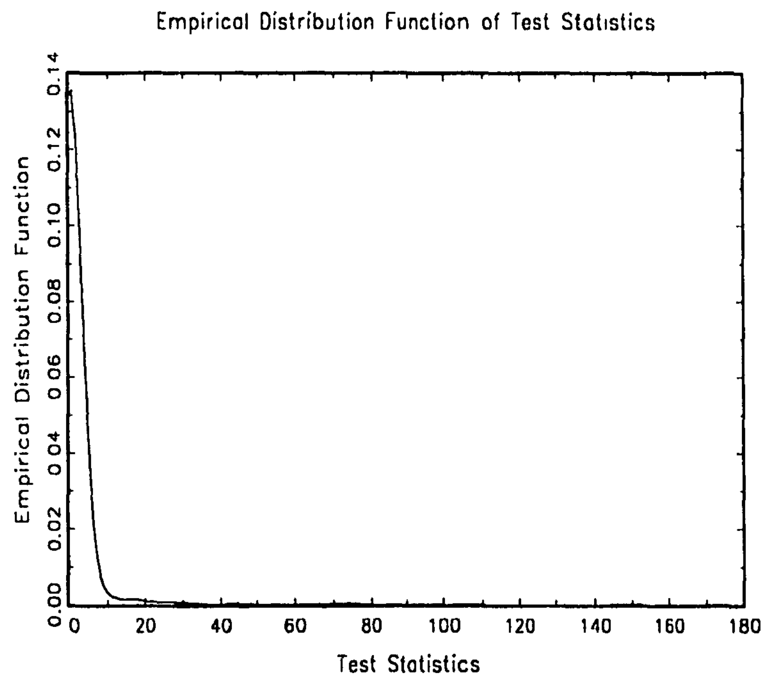


Figure 5.22: The empirical distribution of test-statistics when the null hypothesis is $4D_1 5$

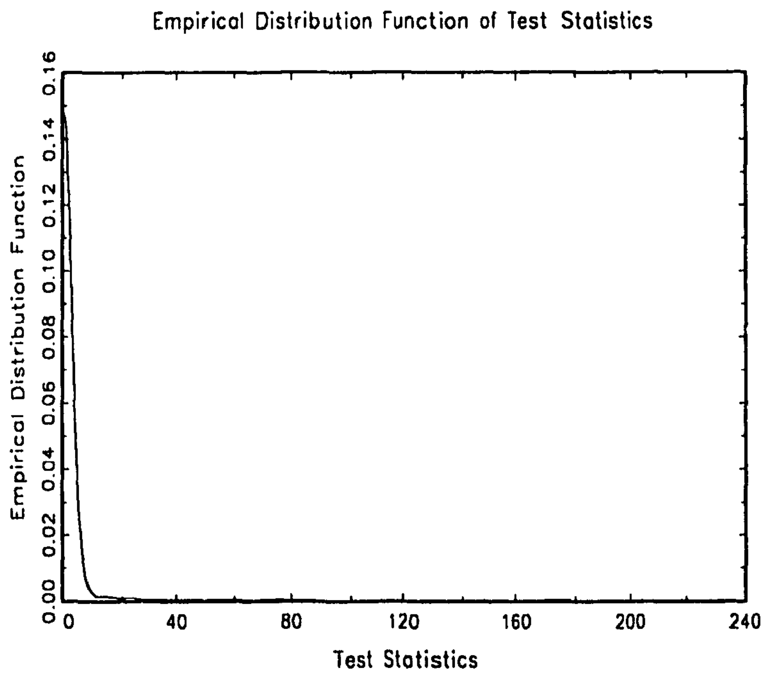


Figure 5.23: The empirical distribution of test-statistics when the null hypothesis is $5D_15$

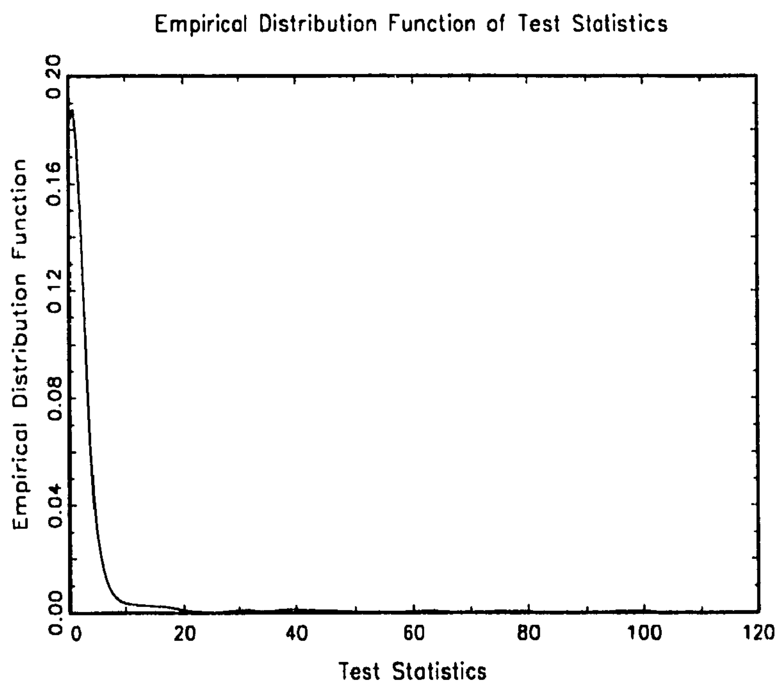


Figure 5.24: The empirical distribution of test-statistics when the null hypothesis is $5D_14$

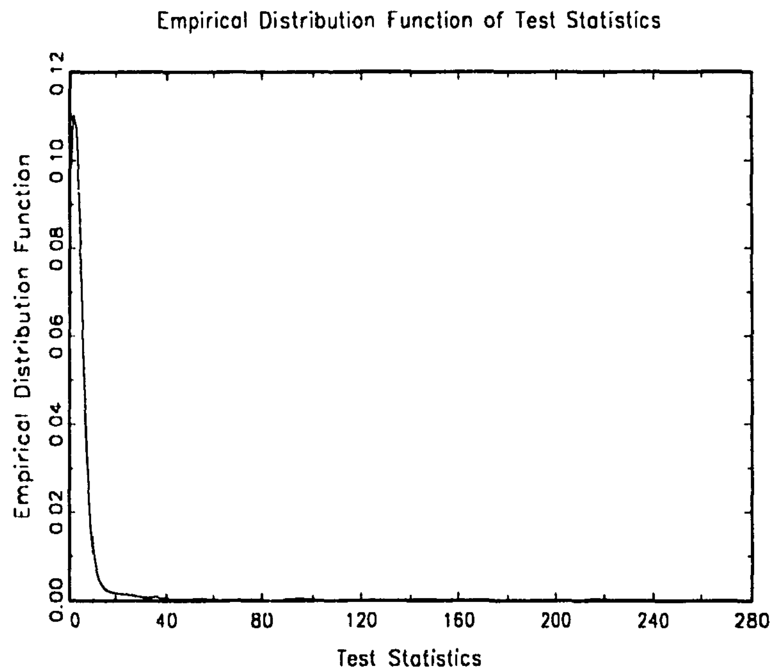


Figure 5.25: The empirical distribution of test-statistics when the null hypothesis is $6D_13$

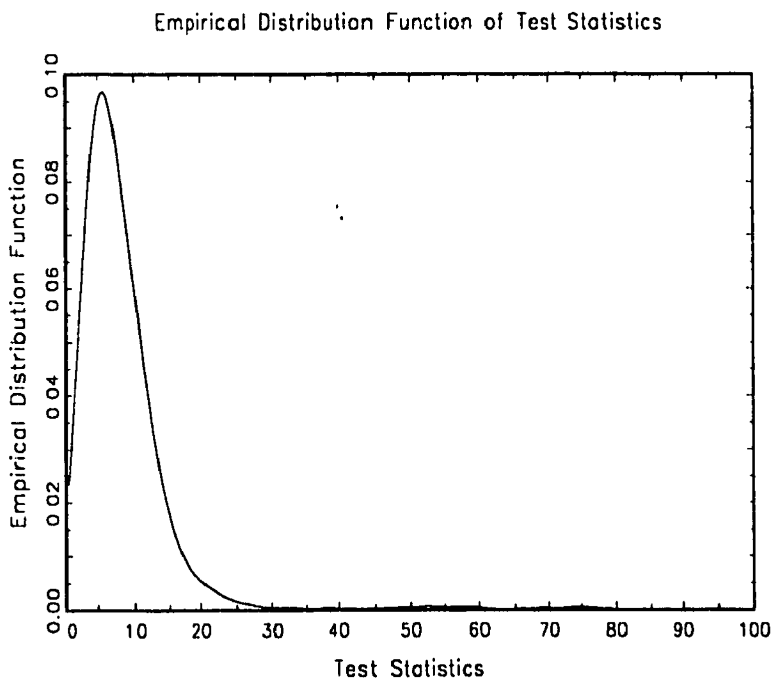


Figure 5.26: The empirical distribution of test-statistics when the null hypothesis is $7D_1 2$

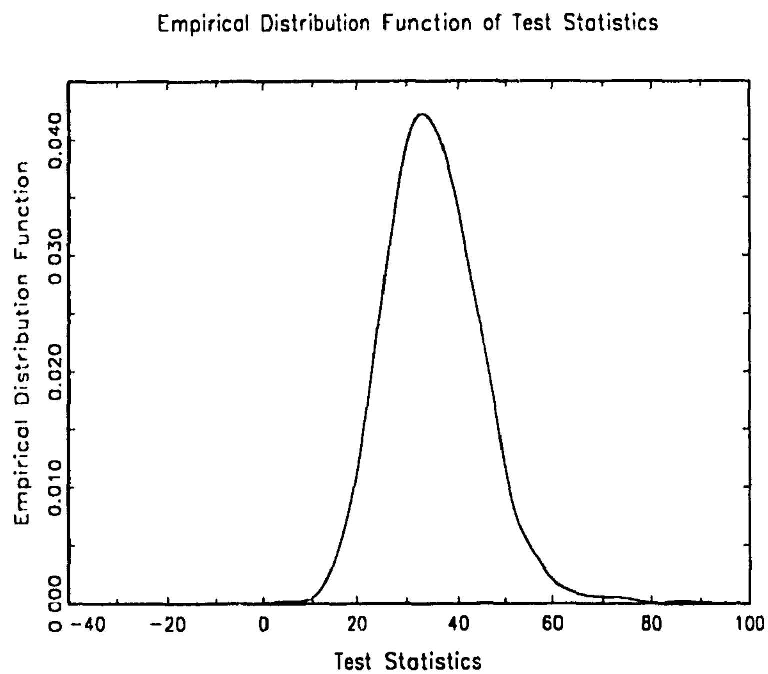


Figure 5.27: The empirical distribution of test-statistics when the null hypothesis is $8D_1 1$

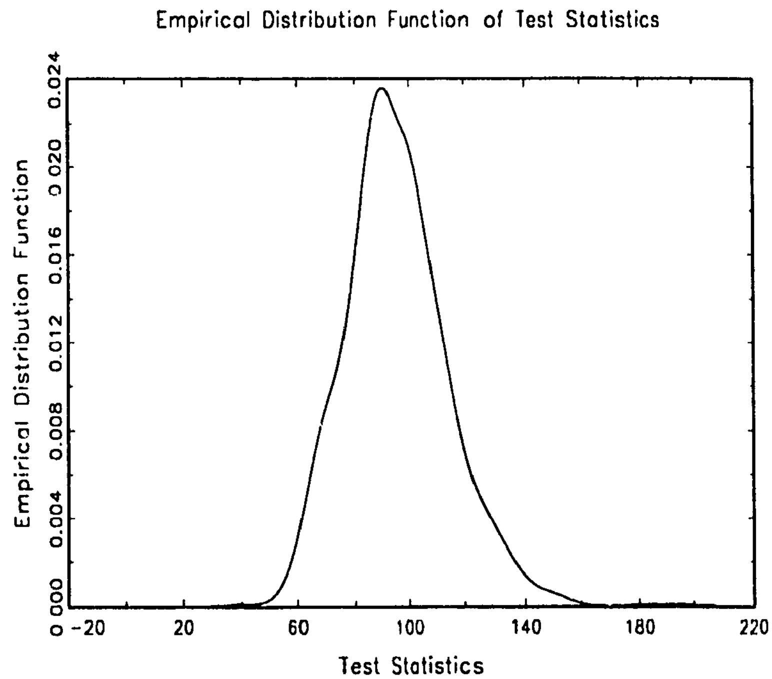


Figure 5.28: The empirical distribution of test-statistics when the null hypothesis is $1D_2 8$

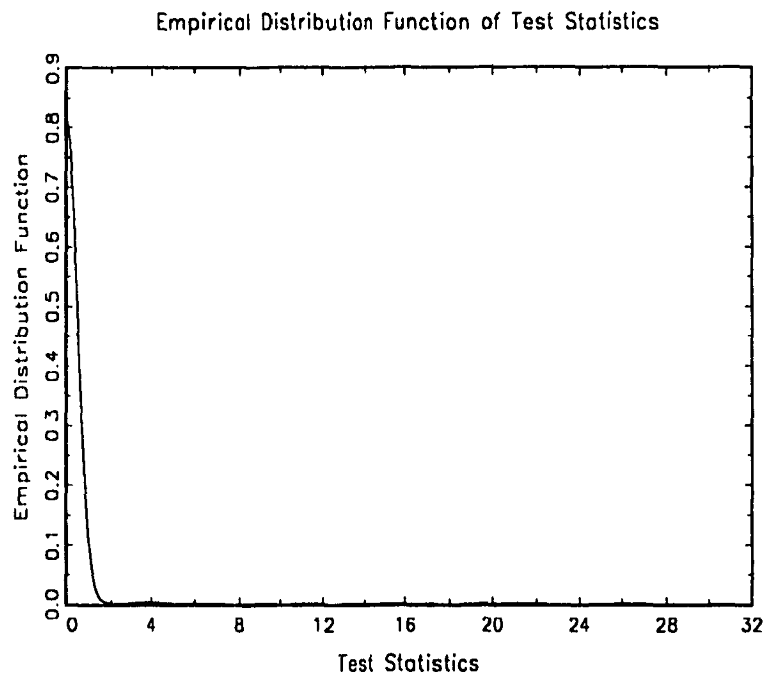


Figure 5.29: The empirical distribution of test-statistics when the null hypothesis is $2D_27$

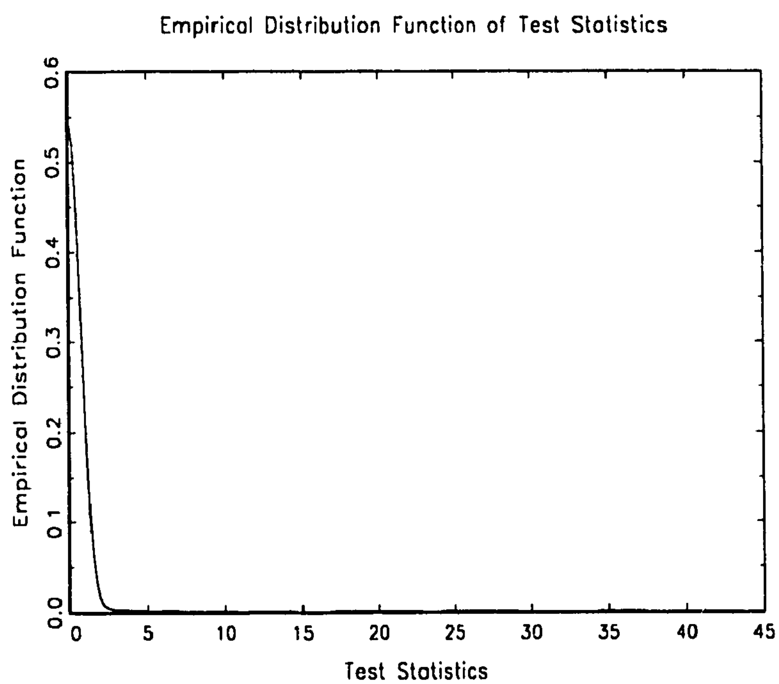


Figure 5.30: The empirical distribution of test-statistics when the null hypothesis is $3D_26$

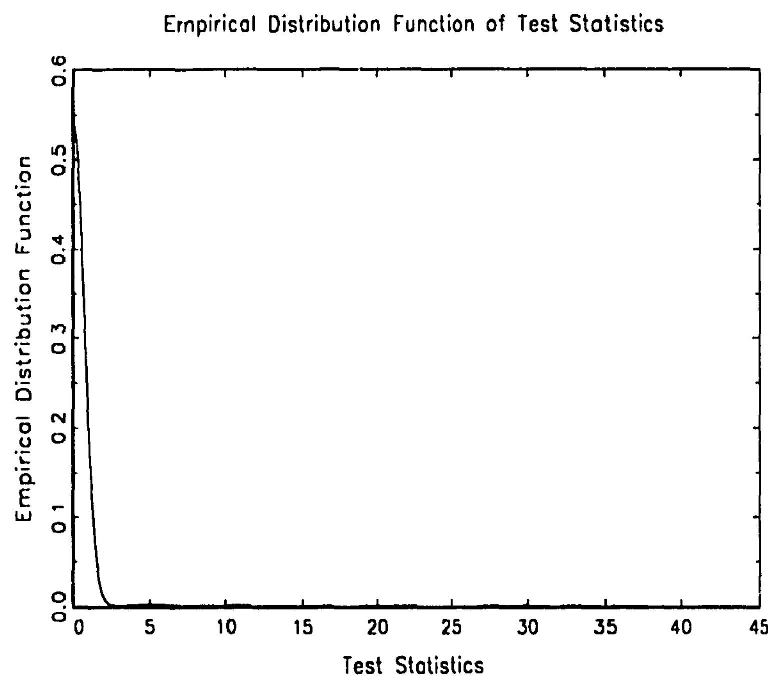


Figure 5.31: The empirical distribution of test-statistics when the null hypothesis is $4D_25$

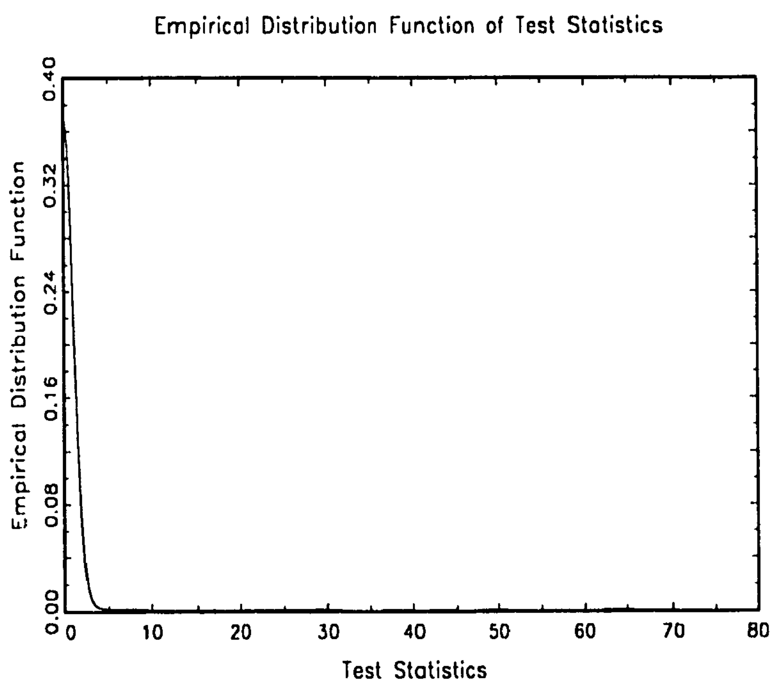


Figure 5.32: The empirical distribution of test-statistics when the null hypothesis is $5D_25$

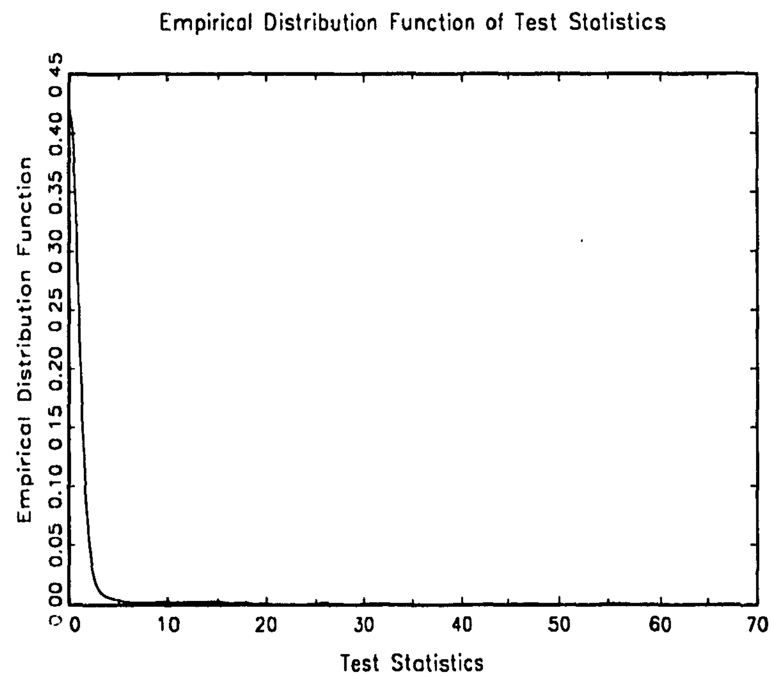


Figure 5.33: The empirical distribution of test-statistics when the null hypothesis is $5D_24$

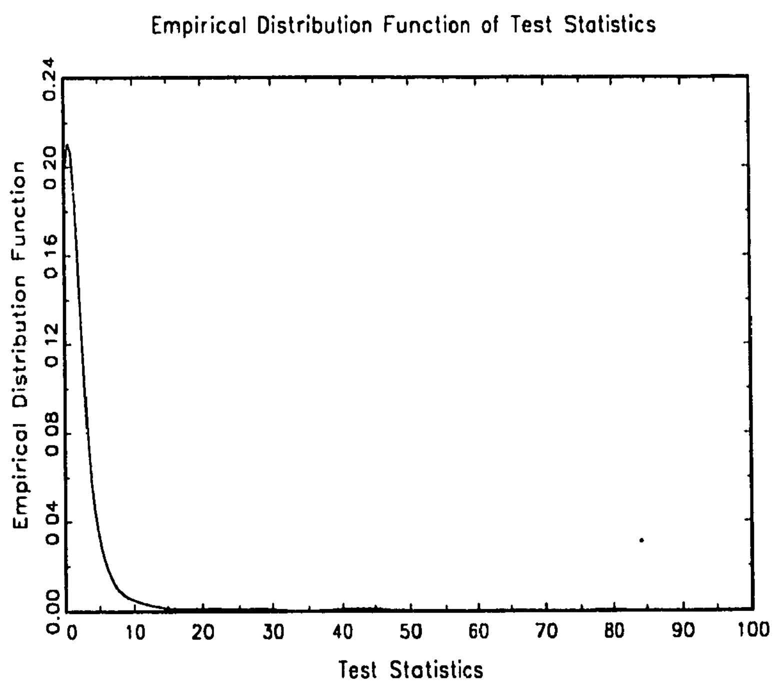


Figure 5.34: The empirical distribution of test-statistics when the null hypothesis is $6D_23$

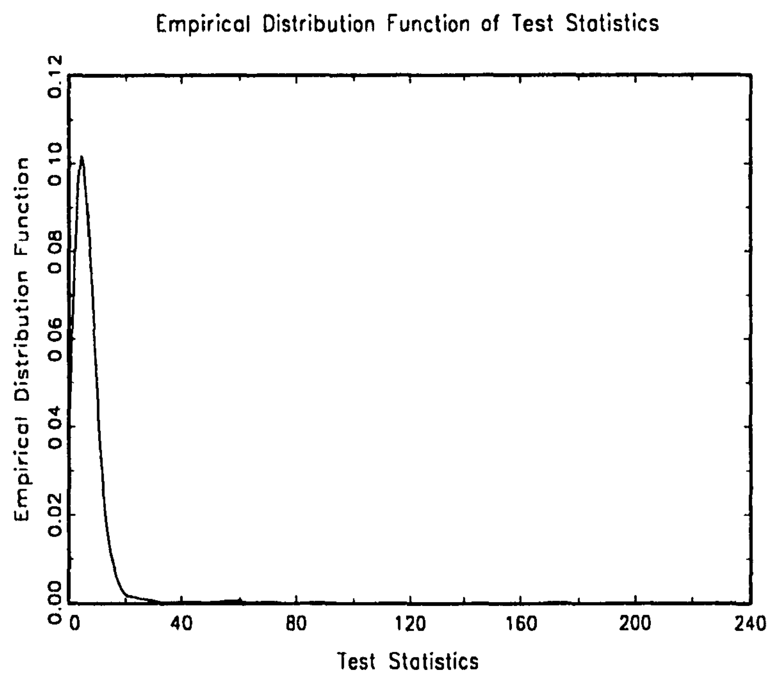


Figure 5.35: The empirical distribution of test statistics when the null hypothesis is $7D_2 2$

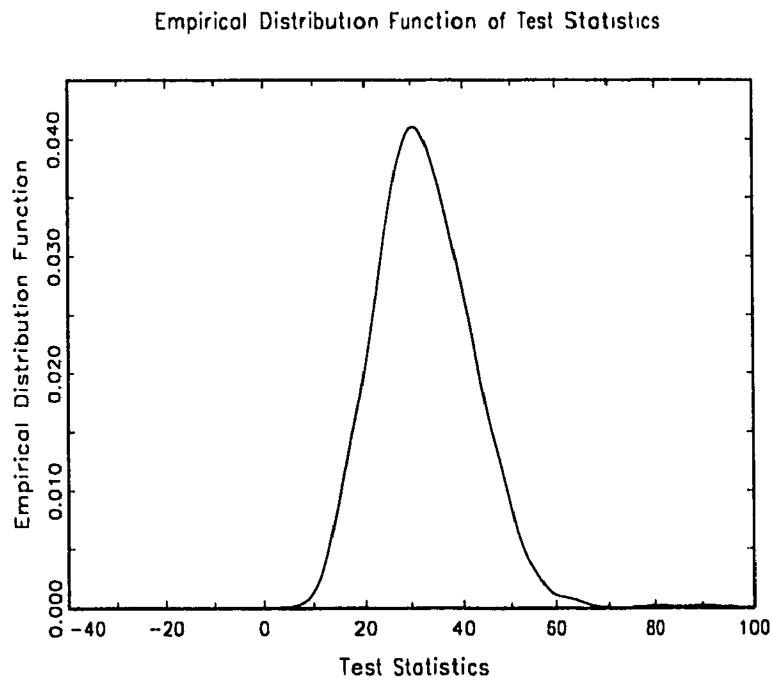
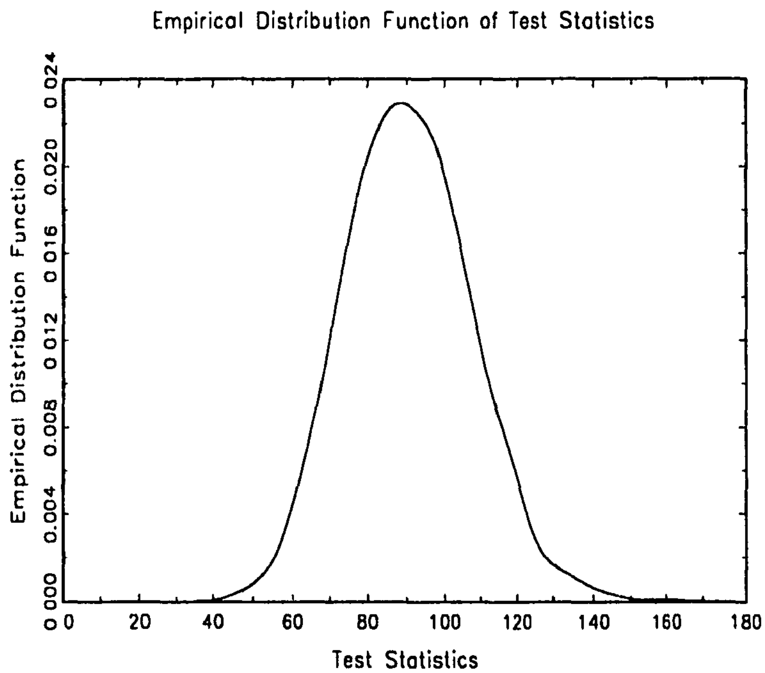


Figure 5.36: The empirical distribution of test-statistics when the null hypothesis is $8D_21$



Chapter 6

Application: Dominance Relations Among T-bill's Holding Period Returns

6.1 Introduction

This chapter illustrates an application of the distribution-free test-statistics for FSD and SSD, c_1 and c_2, given in (4.10) and (4.14). The application is to an analysis of holding period returns of US Treasury bills. This chapter discusses the motivation for such an application, analyzes the basic statistical properties of the returns, implements the dominance tests, and offers concluding remarks.

The existence of term premia in returns for US Treasury bills has long been recognized as an important feature of the term structure [Roll (1970,1971), and Fama (1976)]. While it appears that term premia increase with terms-to-maturity and are statistically significant, it also appears that the variance of holding period returns increases with terms-to-maturity. As a result, it is not clear that investments in long bills would necessarily be preferred to short bills, even though the former offer higher returns on average. This chapter evaluates the economic significance of real term premia in US Treasury bills across terms-to-maturity using the tests for first- and second-degree stochastic dominance advanced in Chapter 4.

In finance, dominance criteria can potentially provide an unambiguous ranking of the desirability of two different assets while placing only general restrictions (i.e. non-satiation and/or risk-aversion) on the preferences of arbitrary investors. Although there are alternative procedures for evaluating the relative attractiveness of different securities, these typically make restrictive assumptions concerning the distribution of returns or the nature of investor preferences. Stochastic dominance is conceptually attractive since it allows comparisons between distributions to be made in a very general way.

The test-statistics, c_1 and c_2, allow time series of individual returns to be weakly dependent, identically distributed random variables,[1] and permit dependence between the random variables whose distributions are to be compared. For the term premia considered in this chapter, the sample correlations between real holding period returns of different terms-to-maturity (one- to six-month) are at least 0.7. One general conclusion from previous research that employs dominance criteria is that first-degree stochastic dominance does not provide discriminatory information concerning the relative rankings of assets such as the one- to sixth-month Treasury bills (Levy 1992). But this conclusion has been derived from procedures that ignore both statistical testing and the tests that will accommodate the dependence among returns. Thus the conclusion is invalid in the sense that its quantitative basis relies on methods which cannot accommodate the data used to reach conclusion. This is not to say that the conclusion is wrong, of course; it may indeed be right—but for the wrong reasons.

[1]In particular, the sequence of returns is assumed to result from ϕ-mixing. Under this assumption, returns may be conditionally heterogeneous but must be unconditionally homogeneous. The tests will be valid for standard *ARCH* processes, for instance.

6.2 Data and Basic Statistical Properties of Data

The yields of 1–6 month Treasury bills for the period of 1952:02–1987:02 are obtained from J. H. McCulloch.[2] These are denoted as $r(t, t+m)$, $m = 1, \ldots, 6$, where t represents the time index and m represents the number of months to maturity. To transform yields to nominal holding period returns, Shiller's (1990) equation is used. The nominal holding period return from t to t' on a bond maturing at time T, $t \leq t' \leq T$, is denoted as $h(t, t', T)$. Shiller's equation depends on the concept of duration. The duration of a bond of term m at time t is defined as:

$$D(m,t) = \frac{\sum_{t_i > t}(t_i - t)s_i e^{-(t_i - t)r(t,t+m)}}{\sum_{t_i > t} s_i e^{-(t_i - t)r(t,t+m)}}, \tag{6.1}$$

where s_i is the coupon payment or principal payment and $r(t, t+m)$ is the yield to maturity from time t to $t+m$. If the bond is a discount bond, then duration equals the term m, i.e.,

$$D(m,t) = m. \tag{6.2}$$

This is because the coupon payments, s_i, are all zero and the last principal payment is one. The nominal holding period return from t to t' on a discount bond maturing at time T, $t \leq t' \leq T$, is defined as:

$$h(t,t',T) = \frac{D(T-t)r(t,T) - [D(T,t) - D(t',t)]r(t',T)}{D(t'-t)}. \tag{6.3}$$

Given that $D(m,t) = m$, the above definition can be simplified to:

$$h(t,t',T) = \frac{(T-t)r(t,T) - (T-t')r(t',T)}{(t'-t)}. \tag{6.4}$$

Equation (6.4) is used to transform yields to nominal holding period returns. Since stochastic dominance tests are to be applied to monthly holding period returns, the

[2]The data are printed in the Appendix of Shiller (1990).

difference of $t' - t$ is always equal to one. The holding period returns are $h(t, t+1, t+m)$, $(m = 1, \ldots, 6)$. The actual transformation causes a loss of one observation in 1987:02. Thus, the data range is 1952:02–1987:01. For simplicity, $h(t, t+1, t+m)$, $m = 1, \ldots, 6$, are denoted as $h(m)$, $m = 1, \ldots, 6$.

In order for the test procedures to be appropriate, the holding period returns should be stationary. However, the augmented Dickey-Fuller tests in Table 6.1 suggest that nominal holding period returns have a unit root, a conclusion that is also supported by Figure 6.1.[3] This nonstationarity is due to the nonstationarity of the rates of inflation. Thus, nominal return series need to be transformed into the real return series.

To transform nominal holding period returns to real holding period returns, the time series of the U.S. consumer price index, CPI_t, is used to compute the time series of the monthly inflation rates,

$$\pi_t = \frac{CPI_t - CPI_{t-1}}{CPI_{t-1}}. \tag{6.5}$$

Given the nominal holding period from time t to $t+1$ on a Treasury bill with m terms-to-maturity, $h(t, t+1, t+m)$, and the inflation rate in the corresponding time period, π_{t+1}, the real holding period return from t to $t+1$ is

$$h^r(t, t+1, t+m) = h(t, t+1, t+m) - \pi_{t+1}, \tag{6.6}$$

for $m = 1, 2, 3, 4, 5, 6$. For simplicity, $h^r(t, t+1, t+m)$, $m = 1, \ldots, 6$, are denoted as $h^r(m)$, $(m = 1, \ldots, 6)$. Real holding period returns are stationary, as verified by the augmented Dickey-Fuller tests in Table 6.2. The graph for real holding period returns are contained in Figure 6.2.

The basic statistical properties of the return series are also examined. Table 6.6 provides additional summary information regarding each of the time series.

[3]The augmented Dickey-Fuller test-statistics are computed by estimating the regression $(y_t - y_{t-1}) = \alpha_0 + \alpha_1 y_{t-1} + \sum_{j=1}^{l} (y_{t-j} - y_{t-j-1}) + \epsilon_t$, and testing the null hypothesis $H_0: \ \alpha_1 = 0$ using the Dickey-Fuller tables. The Akaike information criterion is used to select the number of lags, l, to correct the autoregression in the time series y_t. It turns out that all the test-statistics for properly chosen l's were less than -2.75, the asymptotic critical value at 10%.

Figure 6.3 and Table 6.6 indicate that both the mean of real holding period returns and the variance increase as the term-to-maturity becomes longer. In addition, skewness and kurtosis also increase with term-to-maturity.

One of the important features of the test-statistics advanced in Chapter 4 is that they can properly account for the intricate dependence structure that is a characteristic of most financial data. Practical implementation of these test-statistics requires that the variance-covariance matrix of the differences in sample quantiles be properly estimated; in turn, this requires that the data are in fact weakly dependent, and that some diagnostic information can be used accurately to determine an acceptable moving block size.

Inspection of the sample autocorrelation functions for each of these series (see Table 6.3) indicates that real holding period returns have a non-trivial serial correlation structure, and that the extra complexity associated with the MBB variance-covariance matrix estimate is warranted. However, the autocorrelations die out very slowly. This suggests that real holding period returns may actually be characterized by long- rather than short-range dependence, i.e. real holding period returns may be fractionally integrated (Lo, 1991).

To examine if the time series (the real return series of Treasury bills) are weakly dependent but without a long memory structure, the modified rescaled range test(R/S) developed by Lo (1991) is applied to the series. This test for long-memory is robust to short-range dependence. Assume the time series is denoted as $\{x_1, x_2, \ldots, x_T\}$. The R/S statistic, $R/S_T(l)$, is defined as:

$$R/S_T(l) = \frac{1}{\hat{\sigma}_T(l)} \left[\max_{1 \le k \le T} \sum_{t=1}^{k} (x_t - \bar{x}_T) - \min_{1 \le k \le T} \sum_{t=1}^{k} (x_t - \bar{x}_T) \right], \tag{6.7}$$

where $\bar{x}_T$ is the mean of $\{x_1, x_2, \ldots, x_T\}$ and $\hat{\sigma}_T(l)$ is defined as:

$$\hat{\sigma}_T(l) = \frac{1}{T} \sum_{t=1}^{T} (x_t - \bar{x}_T)^2 + \frac{2}{T} \sum_{t=1}^{l} (1 - \frac{l}{l+1}) \left[\sum_{s=t+1}^{T} (x_s - \bar{x}_T)(x_{s-t} - \bar{x}_T) \right]. \tag{6.8}$$

Under the null hypothesis of no long-memory, the test-statistic has an asymptotic distribution:

$$V_T(l) \equiv \frac{1}{\sqrt{T}} R/S_T(l) \xrightarrow{a} V,$$

where the distribution F_V of V is given by

$$F_V(v) = 1 + 2\sum_{k=1}^{\infty}(1 - 4k^2v^2)e^{-2(kv)^2}. \tag{6.9}$$

Using (6.9), critical values can be computed for any significance level. Lo (1991) provides a table for commonly used significance levels and corresponding critical values [Lo (1991), Table II]. When a test is conducted at the 95 percent level of confidence, the decision rule accepts or rejects according to whether $V_T(l)$ is or is not contained in the interval [0.809, 1.862]; this assigns equal probability to each tail.[4]

As reported in Table 6.4, starting from $l = 40$, the statistic, $V_T(l)$, computed for the real holding period returns of 1-6 month Treasury bills appears to level-off.[5] That implies that there is little evidence that the six series have long-memory and the application of the more general test-statistics, c_1 and c_1, that can accommodate weak dependence within each time series, is appropriate. It also suggests that the tests should use a large value of b.

Table 6.5 contains contemporaneous correlation estimates for the six series. As mentioned previously, the values of the sample correlation coefficients between two real holding period returns ranges from 0.7082 to 0.98181. In addition, it appears that the correlation decreases as the difference in terms-to-maturity increases. When two return series are compared in terms of FSD and SSD, the level of dependence should be duly considered. The MBB estimation introduced in Chapter 4 can accommodate the dependence between the two return series as well as the weak dependence within each series.

[4]One practical issue is the choice of l in (6.7). While Andrews (1991) does provide a data-dependent rule for choosing l, that is still based on an asymptotic mean-squared error criterion, Lo (1991) realises that little is known about how best to pick l in finite samples and reports the statistics for different choices of l. The test is applied to the six return series of Treasury bills at various lags. Note that l should equal $b - 1$.

[5]That is, these statistics fall in the 95% acceptance region [0.809, 1.862]. Please also note that the autocorrelation functions indicate that the the dependence dies out at about lag 50.

6.3 Testing for Stochastic Dominance for T-bill's Holding Period Returns

Figures 6.4, 6.5, and 6.6 are graphs of the empirical quantile functions for the six holding period returns. In general, graphical evidence may provide some useful information regarding the nature of the dominance relationships between returns. For example, Figures 6.4, 6.5, and 6.6, suggest that $h^r(2)$, $h^r(3)$, and $h^r(4)$ may dominate $h^r(1)$; however, it is important to emphasize that tests for stochastic dominance must reflect the fact that distribution and quantile function estimates are subject to sampling errors. Obviously, it is not possible to come to proper statistical conclusions regarding dominance relationships without performing tests. For the null hypotheses of first- and second-degree stochastic dominance at $\alpha = 0.05$, each null hypothesis will be rejected if the test-statistic is greater than 30.841; it will not be rejected if the test-statistic is less than 2.706.[6]

Table 6.7 contains the test results for first-degree stochastic dominance relations among selected pairs of the holding period returns. The moving block bootstrap estimates of the variance-covariance matrix for the difference between two sets of quantile estimates are computed through 200 replications with various block sizes b=10, 20, 30, 40, 50, and 60. The test results are then computed using these estimated MBB variance-covariance matrices. For each pair, dominance relations in both directions are tested. For the reported test-statistics, 20 equally spaced quantiles are selected. The first column of Table 6.7 reports dominance relations under the null hypothesis and test-statistics corresponding to b=10, 20, 30, 40, 50, and 60 are given in columns 2–7, respectively. From Table 6.7, the test-statistics show that the longer-term returns $h^r(2)$, $h^r(3)$, $h^r(4)$, $h^r(5)$, and $h^r(6)$ dominate, in the first-degree, the one-month return ($h^r(1)$), at the 5% level. But as noted earlier, the lower bound of the critical value is very conservative. Thus, there are strong evidence that

[6] If the test-statistic is in between these two critical values and there is a need for a clearer decision rule, the weights in the test-statistic must be calculated.

the long-term returns indeed dominate the very short-term return. As some kind of backup of the above conclusion, it might be useful to consider tests in the opposite direction. From the Table 6.7, it is evident that there is less evidence supporting the null hypothesis that the short-term-end return dominates the longer-term returns in the first-degree (the test-statistics are almost always larger), because the statistics are generally higher than the lower bound of the critical value. But it is not clear that this evidence is statistically significant since the upper bound of the critical value is much higher that the calculated statistics.[7] It could be concluded that the short-term-end return is clearly dominated by the long-term returns, in the sense of first-degree stochastic dominance.

Table 6.8 shows the test results of second-degree stochastic dominance relations among the selected pairs of the holding period returns. These test-statistics reveal that the longer-term returns $h^r(2)$, $h^r(3)$, $h^r(4)$, $h^r(5)$, and $h^r(6)$ dominate, in the second-degree, the one-month return $h^r(1)$. All the null hypotheses can safely be accepted at 5%. Thus, there are strong evidence that longer-term returns indeed dominate the short-term-end return. When considering the tests in the opposite direction, it is found that there is less numerical evidence showing that the opposite dominance relations hold, although it is not clear that the evidence is statistically significant.

It is important to realize that the structure of the data must be accommodated by the assumptions of the test-statistics used to make inferences. For example, the test-statistics for FSD are applied to test whether one-month returns dominate long-term returns, as shown in the lower portion of Table 6.7, $h^r(1)D_1h^r(2)$ has very large values for the calculated test-statistics for various choices of b (viz. 10.726–13.676) while $h^r(1)D_1h^r(6)$ has very small values for the test-statistics for various choices of b (viz. 2.054–2.426). This seems to contradict the observations obtained from the top-left figure of $h^r(2)$ vs. $h^r(1)$ in Figure 6.4, and the bottom-left figure of $h^r(6)$ vs. $h^r(1)$ in Figure 6.5. The empirical quantile curves are much closer to each

[7]This implies that the weights must be computed for the precise critical value.

other for $h^r(2)$ vs. $h^r(1)$ than for $h^r(6)$ vs. $h^r(1)$. Similar observations could be made for the results in the lower portion of Table 6.8.

The question is: What causes such a puzzle? As shown in Chapter 4, the test-statistics need a correctly specified variance-covariance matrix for $\widehat{Q}_X - \widehat{Q}_Y$ (FSD) which involves the estimation of $\frac{1}{T}H\Lambda H'$. The variance-covariance is a function of variance-covariance matrices of $\widehat{Q}_X$ and $\widehat{Q}_Y$ minus two covariance matrices for $\widehat{Q}_X$ and $\widehat{Q}_Y$. It is noted that the dependence between the two returns being compared affects the variance-covariance estimates of $\frac{1}{T}H\Lambda H'$. Generally, the higher the dependence, the smaller the variance-covariance matrix will be. As shown in Table 6.5, for $h^r(1)$ and $h^r(2)$, the estimate of the correlation coefficient is 0.96835, while it is only 0.70824 for $h^r(1)$ and $h^r(6)$. Thus, the variance-covariance matrix of $\widehat{Q}_X - \widehat{Q}_Y$ is smaller for $h^r(1)$ and $h^r(2)$ than for $h^r(1)$ and $h^r(6)$. Because of this, the test-statistic for $h^r(1)D_1h^r(2)$ is higher than that for $h^r(1)D_1h^r(6)$.

The direct interpretation of the above tests is that term premia are not only statistically, but also economically, significant. In an uncertain world, dominating and dominated assets do exist. One important distinction between dominance in a world of certainty and stochastic dominance in a world of uncertainty is that the state probabilities are irrelevant in the former case while being crucial in the latter. In a world of certainty, dominated assets cannot exist or investors could make unlimited arbitrage profits. In a world of uncertainty, stochastically dominated assets may exist, although not as optimal holdings for any investor. Investors might hold a stochastically dominated asset as a part of a portfolio with other assets.[8]

In comparison with Levy and Brooks (1989), it is interesting that our results suggest that at least some longer bonds dominate the shortest bond, in both first- and second-degree. Levy and Brooks ignore the sampling errors associated with the distribution function estimates, yet concluded that no assets are dominated in the first-degree for their entire sample period. Our results, however, indicate that the short bond is dominated, when sampling errors are taken into account. This result

[8]See Ingersoll (1987), p.72–73.

occurs because the sample quantile (or distribution) functions for the shortest and the longer holding period returns cross in the tails of the distributions which implies that no FSD relationship could be found if sampling errors are ignored. However, when sampling errors are properly taken into account, the crossing empirical quantile or distribution functions must be evaluated statistically. More generally, the result that FSD does not appear to be useful in reducing the size of the efficient set [as suggested in Levy (1992)] can be questioned, if sampling errors are duly considered.

6.4 Alternative Procedures For Ranking Assets

As a useful comparison, several other measures besides the mean and standard deviation are considered. These are:

1. Sharpe's index
$$I_S = \frac{(\bar{R}_i - r)}{\sigma_i}, \tag{6.10}$$
where $\bar{R}_i$ is the average return on the i-th investment, r is the risk-free rate, and the risk index is σ_i, the standard deviation of returns [Sharpe (1966)].

2. Treynor's performance index
$$I_T = \frac{(\bar{R}_i - r)}{\beta_i}, \tag{6.11}$$
where β_i is the slope coefficient of the market model [Treynor (1965)].

3. Jensen's excess return α_i, derived from the regression
$$(R_{it} - r_t) = \alpha_i + \beta_i(R_{mt} - r_t) + \varepsilon_{it} \tag{6.12}$$
[Jensen (1968)].

It should be noted that Sharpe's index evaluates the standardized excess returns of an asset or a portfolio. Treynor's performance index is similar to Sharpe's index, but β_i is used. However, Treynor's measure assumes risk aversion and various other assumptions of the capital asset pricing model (CAPM); in particular, riskless borrowing and lending, and normality. Jensen's performance index makes the same assumptions as Treynor's; but it does not necessarily yield the same ranking.

The real series are used to compute the measures. The S&P 500 index is used as the market proxy. The real return is computed from the S&P 500 index. Table 6.9 shows that the three measures for Treasury bills of different holding periods. While 2-month returns are clearly superior according to Sharpe's Index and Treynor's Index, Treynor's index reaches the peak at the 5-month returns.

Mean-variance efficient frontiers are also computed to see if some asset is not in the efficient set. Table 6.10 presents the efficient frontier without the market index, while Table 6.11 illustrates the efficient frontier with the market index. In both cases, no particular asset is excluded from the efficient set. However, it can be observed that the weights for the longer maturities Treasury bills (and the market index) get higher as the requited return of the portfolio is increased.

The differences between these alternative ranking methods and the dominance criteria are that the latter is much more general and place less restriction on the way investors evaluate risky prospects. It is clear that the mean-variance efficient frontiers provide minimum variance portfolios at various desirable level of expected returns.

The Sharp, Treynor, and Jensen all use a risk-free rate as a benchmark. This will automatically exclude the one-month Treasury bill from consideration. Within this smaller set of assets, the two-month Treasury bill gets the highest ranking by the Sharp and Treynor indices. It should also be noted that both Treynor and Jensen measures require similar assumptions as CAPM does, which is rather restrictive. But Jensen's measure behaves so different from Treynor's measure that it judges that five-month Treasury bill to be superior.

When there is a need to identify the least preferred assets without imposing strong assumptions on preferences, dominance tests become useful. First, these tests are less restricted compared to the other ranking measures. For example, Sharp's index, Treynor's performance index, Jensen's excess return, and mean-variance efficient portfolio approach need strong assumptions concerning distributions. Second, the test-statistics are statistically reliable, because the data structures can be properly accommodated by the test-statistics. Third, there is no need to use a risk-free asset in constructing the ranking measures as Sharp's index, Treynor's performance index, and Jensen's excess return do.

6.5 Concluding Remarks

In this chapter, the new distribution-free tests for first- and second-degree stochastic dominance are applied to evaluating the economic significance of real term premia. These test procedures are advantageous because (i) they allow for returns on different assets to be dependent, (ii) they do not restrict sample observations for a particular return to be independent, and (iii) they are distribution-free. These tests were applied to McCulloch's U.S. T-bill data as given in Shiller (1990). The results of first- and second-degree stochastic dominance tests suggest that only the return of the one month Treasury bill is significantly dominated, in the first-degree, by the returns of all the longer term returns in the data set. It is also significantly dominated in the second-degree by the returns of all the longer term returns in the data set. In comparison with existing results, the results in this thesis indicate that it is important to place dominance analysis properly within a proper framework of statistical inference, in the sense that standard results concerning FSD that are derived without accounting for sampling errors may be incorrect.

Table 6.1: The Augmented Dickey-Fuller Tests on Nominal Returns: 1954:02–1987:01

Series	*Test – Statistic t*	*Asymptotic Critical Value at* 10%	*l*
$h(1)$	-1.98	-2.57	5
$h(2)$	-1.89	-2.57	7
$h(3)$	-1.98	-2.57	7
$h(4)$	-2.01	-2.57	8
$h(5)$	-1.90	-2.57	10
$h(6)$	-2.00	-2.57	10

Table 6.2: The Augmented Dickey-Fuller Tests on Real Returns: 1954:02–1987:01

Series	*Test – Statistic t*	*Asymptotic Critical Value at* 10%	*l*
$h^r(1)$	-2.74	-2.57	8
$h^r(2)$	-3.15	-2.57	7
$h^r(3)$	-4.18	-2.57	4
$h^r(4)$	-4.08	-2.57	4
$h^r(5)$	-4.19	-2.57	4
$h^r(6)$	-4.48	-2.57	4

Table 6.3: Autocorrelation Functions for Real Returns: 1954:02–1987:01

	lag 5	*lag* 10	*lag* 15	*lag* 20	*lag* 25	*lag* 30	*lag* 35	*lag* 40	*lag* 45	*lag* 50
$h^r(1)$	.313	.353	.281	.236	.089	.116	.042	.058	-.073	-.051
$h^r(2)$	.323	.338	.314	.241	.124	.022	.076	.061	-.062	-.041
$h^r(3)$	.326	.321	.317	.208	.119	.090	.074	.016	-.048	-.024
$h^r(4)$	.322	.273	.256	.147	.100	-.032	.056	.045	-.045	-.048
$h^r(5)$	.306	.248	.220	.098	.062	.049	.022	-.007	-.049	-.036
$h^r(6)$	.273	.211	.171	.053	.041	-.055	.014	.018	-.045	.023

Table 6.4: R/S Analysis of Real Holding Period Returns of Treasury Bills

$V_T(l)$	1 month	2 month	3 month	4 month	5 month	6 month
$V_T(10)$	3.203	3.048	2.869	2.828	2.656	2.532
$V_T(20)$	1.892	1.768	1.663	1.672	1.586	1.532
$V_T(30)$	1.437	1.321	1.246	1.270	1.216	1.182
$V_T(40)$	1.214	1.102	1.043	1.072	1.033	1.009
$V_T(50)$	1.086	0.976	0.927	0.957	0.930	0.911
$V_T(60)$	1.010	0.898	0.856	0.884	0.866	0.848
$V_T(70)$	0.964	0.852	0.812	0.839	0.827	0.810
$V_T(80)$	0.942	0.829	0.790	0.814	0.806	0.790
$V_T(90)$	0.942	0.827	0.788	0.807	0.802	0.784
$V_T(100)$	0.961	0.843	0.802	0.817	0.813	0.795

Table 6.5: Correlation Matrix for Real Returns: 1954:02–1987:01

	$h^r(1)$	$h^r(2)$	$h^r(3)$	$h^r(4)$	$h^r(5)$	$h^r(6)$
$h^r(1)$	1.00000					
$h^r(2)$	0.96835	1.00000				
$h^r(3)$	0.90843	0.97070	1.00000			
$h^r(4)$	0.84126	0.91277	0.96482	1.00000		
$h^r(5)$	0.77704	0.85711	0.93046	0.98151	1.00000	
$h^r(6)$	0.70827	0.79990	0.88763	0.95484	0.99064	1.00000

Table 6.6: The Basic Statistics of Real Returns: 1954:02–1987:01

Series	*Mean*	*Variance*	*Skewness*	*Kurtosis*
$h^r(1)$	0.06780	0.08100	0.06921	1.90701
$h^r(2)$	0.11238	0.08584	0.14013	1.78988
$h^r(3)$	0.13600	0.10052	0.07462	2.00694
$h^r(4)$	0.15140	0.11646	0.32890	1.86587
$h^r(5)$	0.16031	0.13935	0.50792	2.58086
$h^r(6)$	0.17006	0.16877	0.75224	3.68112

Table 6.7: Test-Statistics of First-Degree Stochastic Dominance Using MBB of Different Block Sizes (b)

H_0	$b = 10$	$b = 20$	$b = 30$	$b = 40$	$b = 50$	$b = 60$
$h^r(2)D_1h^r(1)$	0.000	0.000	0.000	0.000	0.000	0.000
$h^r(3)D_1h^r(1)$	0.052	0.055	0.051	0.049	0.054	0.064
$h^r(4)D_1h^r(1)$	0.000	0.000	0.000	0.000	0.000	0.000
$h^r(5)D_1h^r(1)$	0.004	0.005	0.004	0.004	0.005	0.004
$h^r(6)D_1h^r(1)$	0.101	0.093	0.107	0.107	0.118	0.119
$h^r(1)D_1h^r(2)$	13.676	13.633	12.564	11.069	10.726	10.930
$h^r(1)D_1h^r(3)$	3.824	3.687	4.370	5.178	6.133	4.802
$h^r(1)D_1h^r(4)$	4.454	4.713	4.761	4.583	5.181	4.992
$h^r(1)D_1h^r(5)$	3.272	3.441	3.150	3.715	3.405	3.826
$h^r(1)D_1h^r(6)$	2.054	2.304	2.268	2.426	2.361	2.419

Note: The sample points selected, K, is 20. The number of samples in the moving block bootstrap is 200. At $\alpha = 0.05$, H_0, under which either $Q_Y(P) - Q_X(P) \geq 0\ \forall P \in [0,1]$ or $Q_X(P) - Q_Y(P) \geq 0\ \forall P \in [0,1]$, will be rejected if the test-statistic is greater than 30.841; it will not be rejected if the test-statistic is less than 2.706.

Table 6.8: Test-Statistics of Second-Degree Stochastic Dominance Using MBB of Different Block Sizes (b)

H_0	$b = 10$	$b = 20$	$b = 30$	$b = 40$	$b = 50$	$b = 60$
$h^r(2)D_2h^r(1)$	0.000	0.000	0.000	0.000	0.000	0.000
$h^r(3)D_2h^r(1)$	0.045	0.052	0.054	0.051	0.057	0.057
$h^r(4)D_2h^r(1)$	0.000	0.000	0.000	0.000	0.000	0.000
$h^r(5)D_2h^r(1)$	0.000	0.000	0.000	0.000	0.000	0.000
$h^r(6)D_2h^r(1)$	0.114	0.109	0.109	0.091	0.126	0.126
$h^r(1)D_2h^r(2)$	12.067	9.342	10.087	8.719	6.360	7.707
$h^r(1)D_2h^r(3)$	2.880	2.369	3.161	2.286	2.742	2.476
$h^r(1)D_2h^r(4)$	2.778	2.810	2.888	2.992	2.727	2.788
$h^r(1)D_2h^r(5)$	1.558	1.493	1.381	1.584	1.855	1.494
$h^r(1)D_2h^r(6)$	0.693	0.710	0.752	0.818	0.756	0.792

Note: The sample points selected, K, is 20. The number of samples in the moving block bootstrap is 200. At $\alpha = 0.05$, H_0, under which either $\Phi_Y(P) - \Phi_X(P) \geq 0 \ \forall P \in [0,1]$ or $\Phi_X(P) - \Phi_Y(P) \geq 0 \ \forall P \in [0,1]$, will be rejected if the test-statistic is greater than 30.841; it will not be rejected if the test-statistic is less than 2.706.

Table 6.9: Performance Measures

Month	Sharpe	Treynor	Jensen
1	0.0000	0.0000	0.0000
2	0.6096*	-0.0251*	0.0124
3	0.5144	-0.0323	0.0199
4	0.4531	-0.0358	0.0267
5	0.3936	-0.0353	0.0302*
6	0.3525	-0.0347	0.0293

Note:*—Peak.

Table 6.10: Efficient Frontier Without S & P 500 Index

Mean	Std. Dev.	1 month	2 month	3 month	4 month	5 month	6 month
0.42	0.25	1.00	0.00	0.00	0.00	0.00	0.00
0.42	0.25	0.92	0.01	0.04	0.00	0.00	0.00
0.43	0.25	0.87	0.01	0.02	0.01	0.06	0.00
0.43	0.26	0.69	0.27	0.00	0.00	0.01	0.00
0.44	0.26	0.60	0.28	0.08	0.00	0.01	0.00
0.44	0.26	0.59	0.19	0.09	0.05	0.03	0.02
0.45	0.27	0.53	0.20	0.10	0.06	0.05	0.03
0.45	0.27	0.19	0.79	0.00	0.00	0.00	0.00
0.46	0.28	0.54	0.00	0.00	0.13	0.29	0.02
0.46	0.28	0.35	0.18	0.17	0.12	0.08	0.07
0.47	0.29	0.33	0.29	0.01	0.00	0.00	0.36
0.47	0.31	0.41	0.02	0.00	0.01	0.13	0.39
0.48	0.30	0.00	0.44	0.38	0.09	0.03	0.04
0.48	0.30	0.01	0.06	0.90	0.00	0.00	0.01
0.49	0.32	0.14	0.08	0.25	0.05	0.16	0.29
0.49	0.33	0.00	0.32	0.16	0.00	0.03	0.46
0.50	0.33	0.01	0.00	0.37	0.00	0.60	0.00
0.50	0.35	0.00	0.00	0.24	0.18	0.35	0.21
0.51	0.37	0.02	0.00	0.00	0.31	0.16	0.48
0.51	0.38	0.00	0.00	0.00	0.01	0.49	0.49

Table 6.11: Efficient Frontier With S & P 500 Index

Mean	Std. Dev.	1 month	2 month	3 month	4 month	5 month	6 month	Index
0.42	0.25	1.00	0.00	0.00	0.00	0.00	0.00	0.00
0.43	0.25	0.83	0.01	0.13	0.00	0.00	0.00	0.00
0.44	0.26	0.70	0.02	0.19	0.00	0.00	0.06	0.00
0.45	0.27	0.50	0.19	0.11	0.07	0.06	0.04	0.00
0.46	0.28	0.41	0.07	0.15	0.22	0.10	0.02	0.00
0.47	0.28	0.30	0.18	0.14	0.12	0.10	0.09	0.03
0.48	0.31	0.09	0.00	0.58	0.18	0.12	0.00	0.00
0.50	0.44	0.18	0.20	0.00	0.27	0.20	0.00	0.12
0.51	0.36	0.02	0.00	0.11	0.24	0.19	0.42	0.00
0.52	0.84	0.21	0.07	0.00	0.15	0.24	0.04	0.25
0.53	0.66	0.13	0.00	0.00	0.00	0.00	0.66	0.19
0.54	0.62	0.00	0.00	0.00	0.00	0.04	0.77	0.17
0.55	1.27	0.00	0.00	0.32	0.17	0.02	0.08	0.38
0.56	1.63	0.01	0.16	0.13	0.00	0.10	0.08	0.49
0.57	1.93	0.00	0.22	0.00	0.07	0.11	0.00	0.57
0.59	2.08	0.00	0.05	0.15	0.04	0.04	0.07	0.62
0.60	2.40	0.07	0.01	0.00	0.00	0.19	0.00	0.71
0.61	2.50	0.00	0.00	0.00	0.00	0.21	0.02	0.74
0.62	2.81	0.00	0.00	0.00	0.08	0.05	0.01	0.83
0.63	3.08	0.00	0.00	0.00	0.00	0.05	0.01	0.91

Figure 6.1: The Nominal Holding Period Returns: 1954:02–1987:01

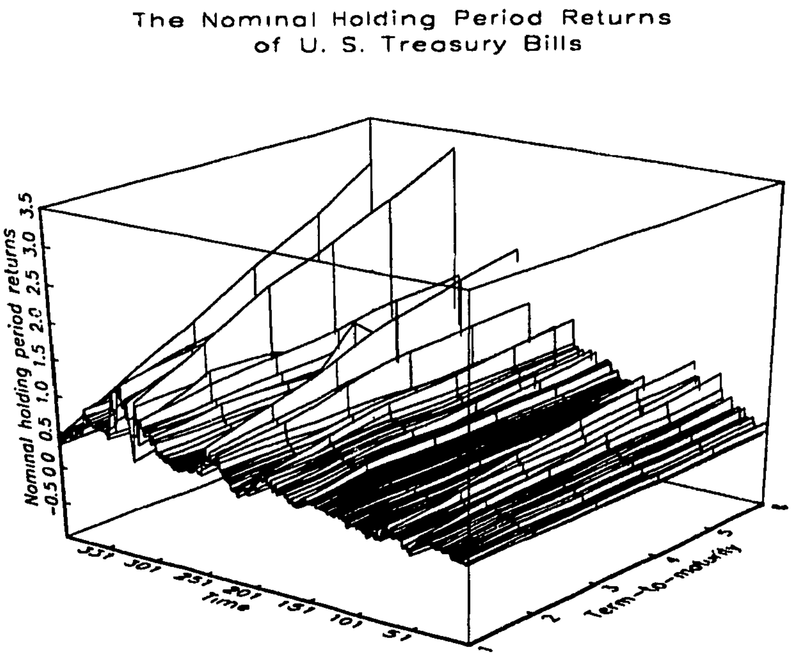


Figure 6.2: The Real Holding Period Returns: 1954:02–1987:01

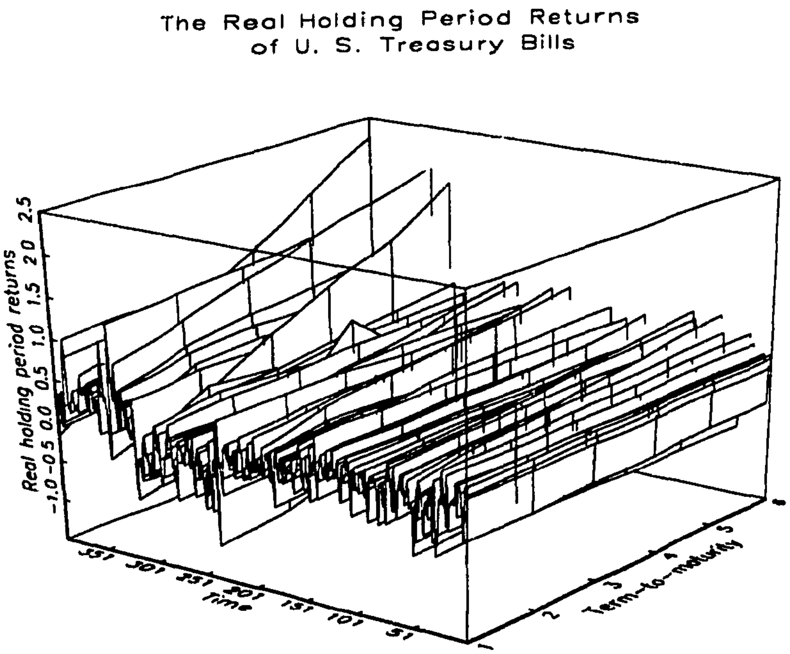


Figure 6.3: The Relation between Mean and Variance across Terms-To-Maturity

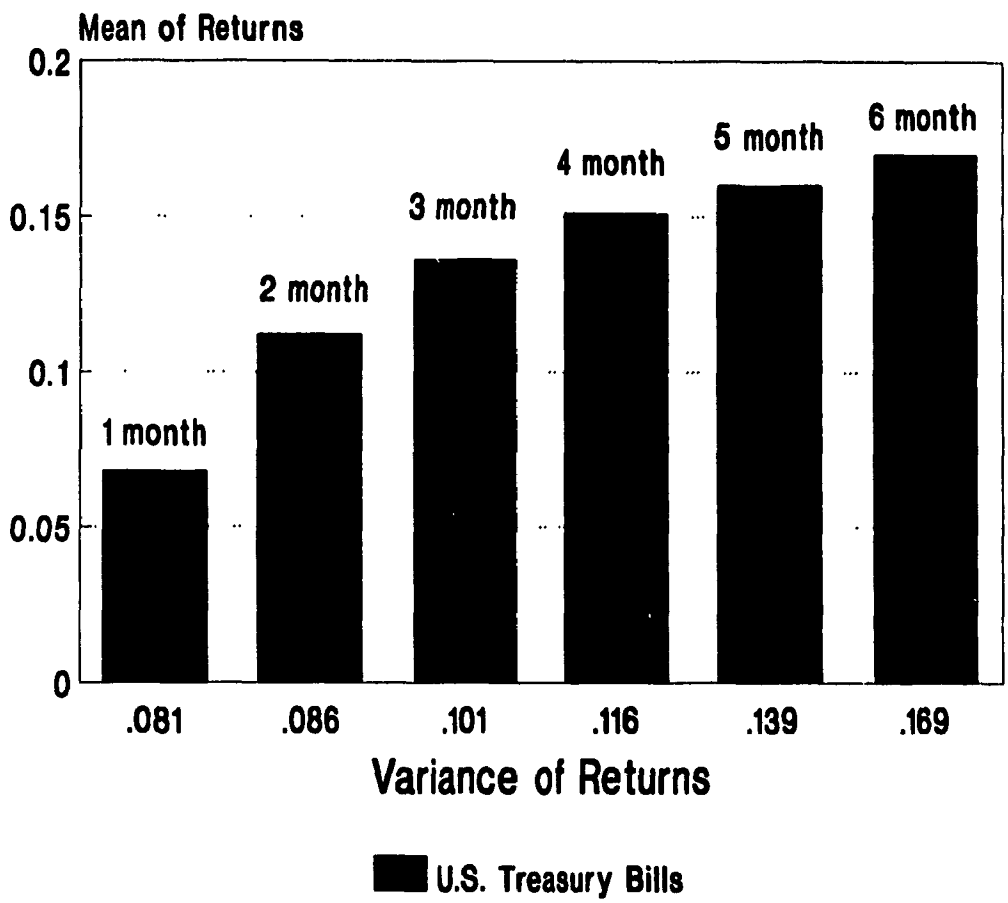


Figure 6.4: Empirical Quantile Functions

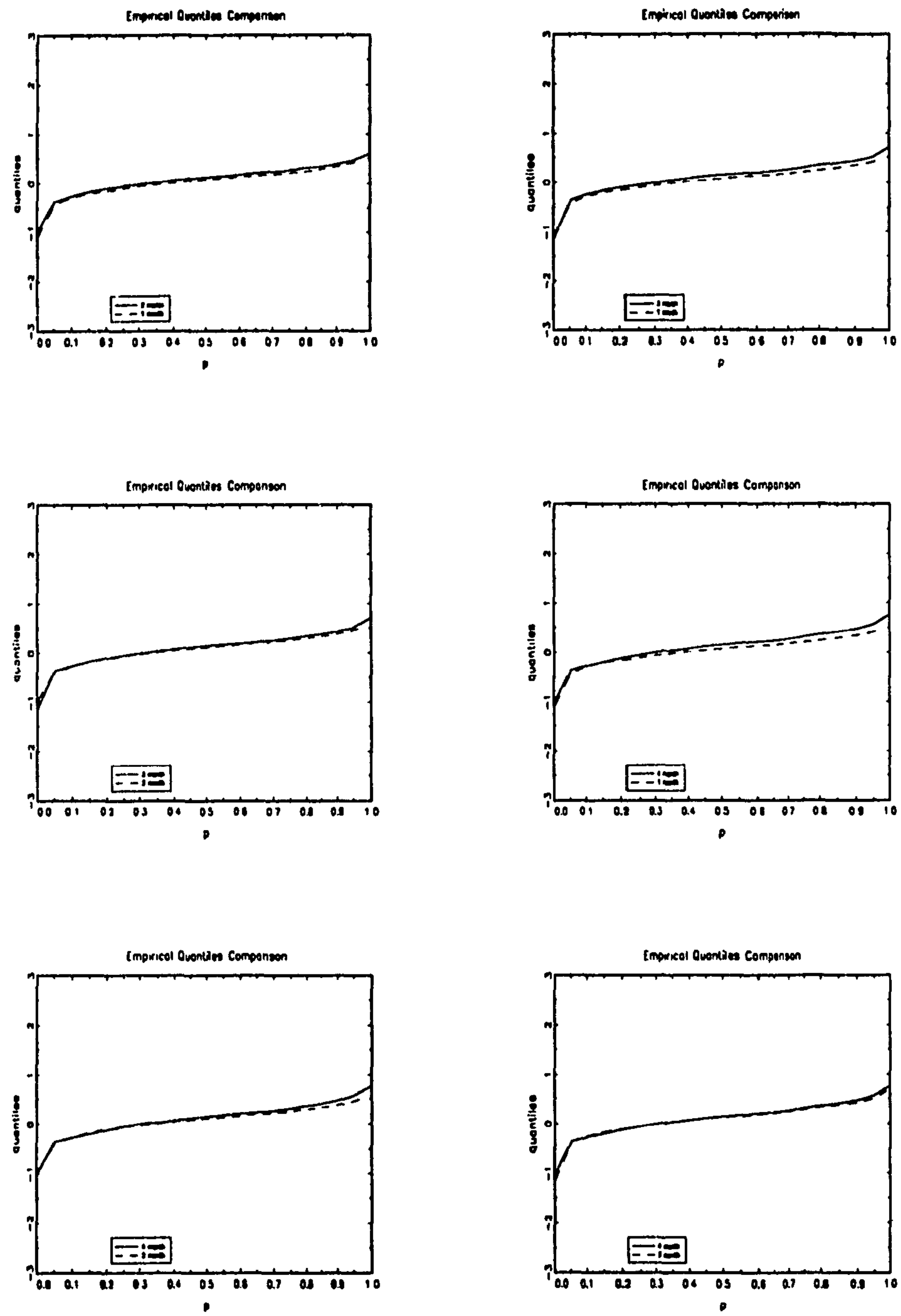


Figure 6.5: Empirical Quantile Functions (Continued)

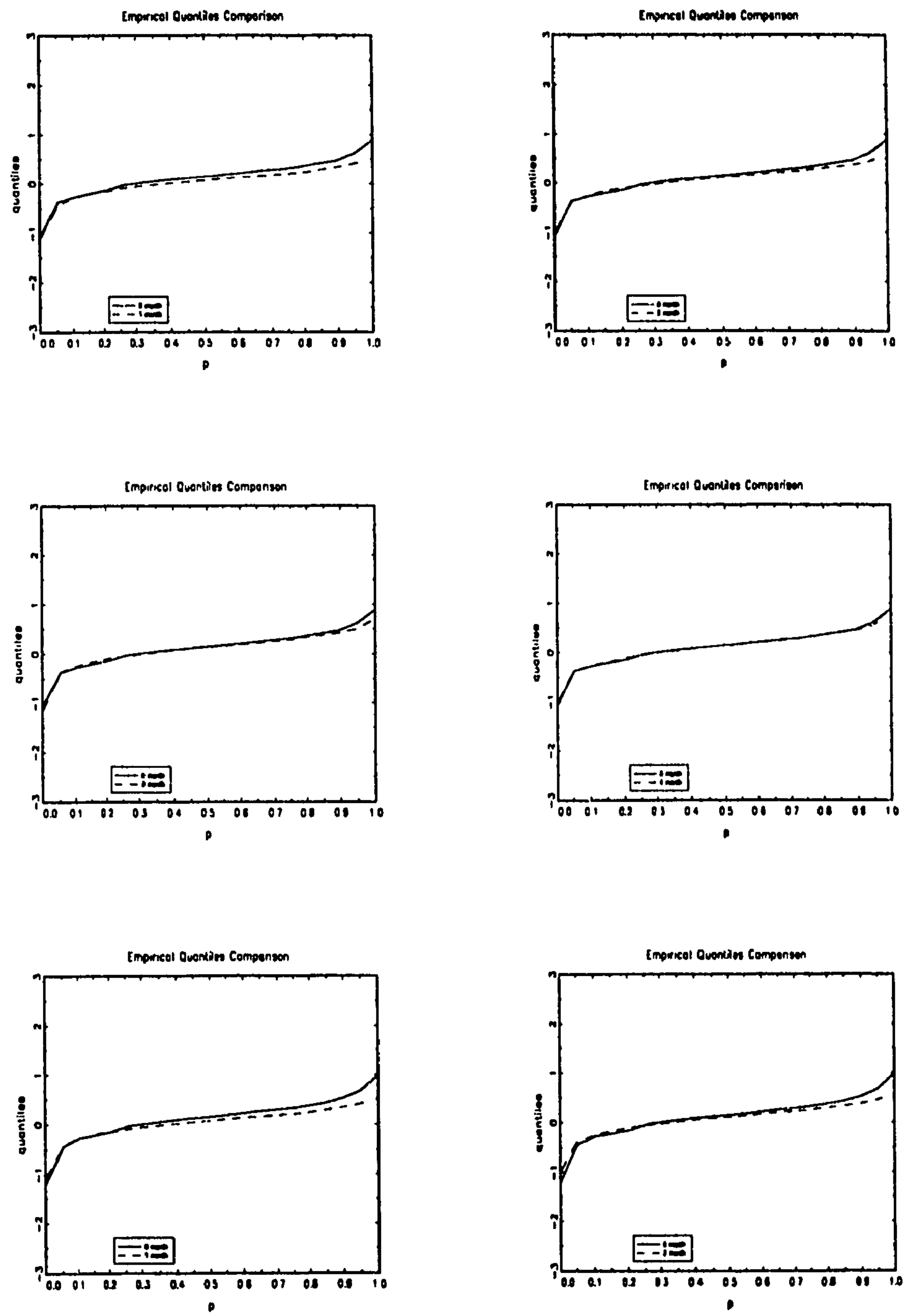
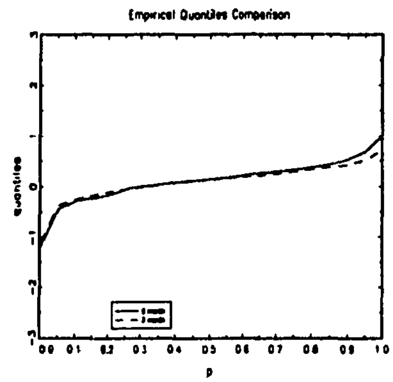
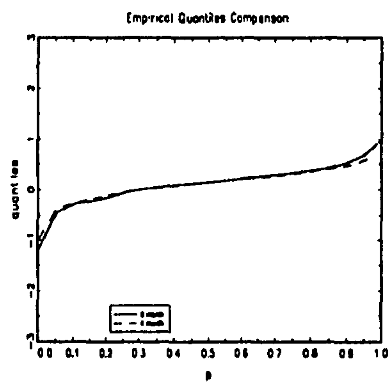
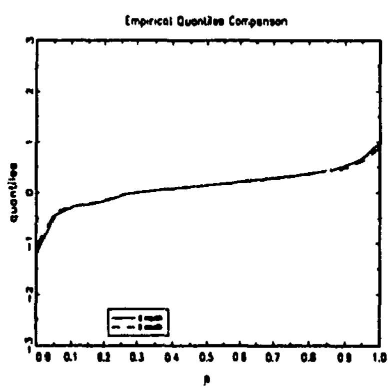


Figure 6.6: Empirical Quantile Functions (Continued)







Chapter 7

Summary and Discussion

Stochastic dominance is an important concept in economics and finance. For example, it is frequently seen that income distributions are compared using stochastic dominance criteria, and that asset returns are ranked according to stochastic dominance criteria in finance. The theoretical criteria of stochastic dominance do not of themselves provide tests for stochastic dominance but merely a foundation for such tests.

To make the criteria testable, it is necessary to establish test procedures. Unfortunately, existing algorithms and test procedures are very restrictive. Some eschew sampling theory; some are restricted to a special class of known parametric distribution functions; some have no clear relation to dominance characteristics; and some have no asymptotic justification.

This thesis advances a new test procedure for first- and second-order stochastic dominance, and its specific forms under various data structures. In Chapter 4, under Assumptions 1, and 2, Theorems 3 and 4 give the tests for first- and second-order stochastic dominance, respectively. c_1^d and c_2^d, given in Corollaries 1 and 2, are the FSD and SSD test-statistics under Assumptions 2 and 3. c_1^i and c_2^i, given in Corollaries 3 and 4, are the FSD and SSD test-statistics under Assumptions 2 and 4.

Monte Carlo simulation results suggest that the test statistics are powerful in the following sense. When one data generating process strictly dominates another, the empirical size of the test is generally smaller than its nominal theoretical size. On the other hand, if the data generating processes do not follow a dominance relation, in whatever degree, under the null hypothesis, the power of the test is generally very large. In other words, when one data generating process fails to dominate another, the empirical power of the test is very close to one. When two data generating processes are similar, the test generally has lower power as one would expect.

The empirical work has demonstrated how to apply the tests when the data can be characterized by a mixing process. In the economic evaluation of the holding period returns of U.S. Treasury bills, it is clear that previous research ignores sampling errors and hence its conclusions do not have a sound scientific basis. Using the tests proposed in this thesis, it is found that the returns of the one-month Treasury bill are clearly dominated by the returns of the longer-term Treasury bills. While the existence of the dominated asset is itself an interesting issue worth exploring, the dominance tests proposed are revealed to be useful in identifying its existence.

The thesis identifies a research area in stochastic dominance, and advances a useful test procedure and two important extensions. The finite sample properties are now understood through the Monte Carlo simulation. An application also serves as a useful demonstration. Several issues are left for future research. These are:

1. The theory and concept of the third-degree stochastic dominance are still debated and in a process of further development. The test issue related to the third-degree stochastic dominance should be a focus in future research.

2. Income distribution analysis is also a dynamic research area. The proposed tests are readily applicable to the income survey data over time and across countries. This task is left for the future.

3. Stochastic dominance tests can be further extended to conditional distribution and quantile functions, which may provide links between dominance relations

and the variables in the relevant information sets. This extension will undoubtedly provide a richer analysis of data in hand.

4. It is noted that when the data generating processes are similar, the test statistics are often in the inconclusive region, the critical value must then be determined by a numerical method (i.e., simulating the weights used in the distribution of the test statistics). This topic has not been pursued and is left for the future.

References

Anderson, G. (1994). "Nonparametric Tests of Stochastic Dominance in Income Distribution," Mimeo, Department of Economics, University of Toronto.

Ait-Sahalia, Y. (1992). "The Delta and Bootstrap Methods for Nonlinear Functionals of Nonparametric Kernel Estimators Based on Dependent Multivariate Data," Mimeo, Department of Economics, Massachusetts Institute of Technology, Cambridge, MA, 02139.

Andrews, D. W. K. (1983). "Heteroskedasticity and Autocorrelation Consistent Covariance Matrix Estimation," *Econometrica*, Vol. 59, 817–858.

Bahadur, R. R. (1966). "A Note on Quantiles in Large Samples," *Annuals of Mathematical Statistics*, 37, 577–580.

Bawa, V. S. (1985). "Optimal Rules for Ordering Uncertain Prospects," *Journal of Financial Economics*, 2, 95–121.

Beach, C. M. and R. Davidson (1983). "Distribution-Free Statistical Inference with Lorenz Curves and Income Shares," *Review of Economic Studies*, 723–735.

Beach, C. M., and J. Richmond (1985). "Joint Confidence Intervals for Income Shares and Lorenz Curves," *International Economic Review*, 26, 439–450.

Bishop, J. A., J. P. Formby, and P. D. Thistle (1989). "Statistical Inference, Income Distributions, and Social Welfare," *Research on Economic Inequality*, 1, 49–82.

Bishop, J. A., S. Chakraborti, and P. D. Thistle (1988). "Relative Inequality, Absolute Inequality, and Welfare: Some Statistical Tests," *Economic Letters*, 26, 291–294.

Bishop, J. A., S. Chakraborti, and P. D. Thistle (1989). "Asymptotically Distribution-Free Statistical Inference for Generalized Lorenz Curves," *Review of Economics and Statistics*, 71, 725–727.

Bollerslev, T. (1987). "A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return," *Review of Economics and Statistics*, 69, 542–546.

Boudoukh, J., M. Richardson, and T. Smith (1993). "Is the *Ex Ante* Risk Premium Always Positive?" *Journal of Financial Economics*, 34, 387–408.

Boyd, D. W., and J. M. Steele (1978). "Lower Bounds for Nonparametric Density Estimation Rates," *Annals of Statistics*, 6, 932–934.

Bradley, R. C. (1986). "Basic Properties of Strong Mixing Coefficients," in E. Eberlein and M. S. Taqqu (ed.), *Dependence in Probability and Statistics*, Birkhauser, 165–192.

David, H. A. (1981). *Order Statistics*, John Wiley and Sons.

Davidson, R. and J. G. MacKinnon (1993). *Estimation and Inference in Econometrics*, Chapter 21 Monte Carlo Experiments, 731–768.

Deshpande, J. V. and H. Singh (1985). "Testing for Second Order Stochastic Dominance," *Communications in Statistics: Theory and Methods*, 14, 887–897.

Devroye, L. (1983). "The Equivalence of Weak, Strong, and Complete Convergence in L_1 for Kernel Density Estimates," *Annals of Statistics*, 11, 896–904.

Devroye, L., and L. Gyorfi (1985). *Nonparametric Density Estimation: The L_1 View*, Wiley.

Durbin, J. (1973). *Distribution Theory for Tests Based on the Sample Distribution Function*, SIAM, Philadelphia.

Durbin, J. (1985). "The First Passage Density of a Continuous Gaussian Process to a General Boundary," *Journal of Applied Probability*, 22, 99–122.

Edginton, E. S. (1980). *Randomization Tests*, Marcel Dekker.

Efron, B. (1982). *The Jackknife, the Bootstrap, and Other Resampling Plans*, SIAM.

Efron, B. and R. Tibshirani (1986). "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," *Statistical Science*, 1, 54–77.

Fama, E. (1976a). "Inflation Uncertainty and Expected Returns on Treasury Bills," *Journal of Political Economy*, 84, 427–488.

Fama, E. (1976b). "Forward Rates as Predictors of Future Spot Rates," *Journal of Financial Economics*, 3, 361–377.

Fama, E. (1984a). "The Information in the Term Structure," *Journal of Financial Economics*, 13, 509–528.

Fama, E. (1984b). "Term Premium in Bond Returns," *Journal of Financial Economics*, 13, 529–546

Farrell, R. H. (1972). "On the Best Obtainable Asymptotic Rates of Convergence in Estimation of a Density Function at a Point," *Annals of Mathematical Statistics*, 43, 170–180.

Fernholz, L. T. (1983). *Von Mises Calculus for Statistical Functionals*, Lecture Notes in Statistics 19, Springer-Verlag.

Fryer, M. J. (1976). "Some Errors Associated with the Nonparametric Estimation of Density Functions," *Journal of the Institute of Mathematics and Its Applications*, 18, 371–380.

Gnedenko, B. and V. Koroyuk (1961). "On the Maximum Discrepancy between Two Empirical Distributions," *Selected Translations in Mathematical Statistics and Probability*, 1, 13–16.

Hadar, J. and W. R. Russell (1969). "Rules for Ordering Uncertain Prospects," *American Economic Review*, 59, 25–34.

Hanoch, G. and H. Levy (1969). "The Efficiency Analysis of Choices Involving Risk," *Review of Economic Studies*, 36, 335–346.

Hendry, D. (1984). "Chapter 16 Monte Carlo Experimentation in Econometrics," in: Z. Griliches and M. D. Intriligator, eds., *Handbook of Econometrics*, Vol. 2, 937–976.

Hochberg, Y. (1974). "Some Generalizations of the T-method in Simultaneous Inference," *Journal of Multivariate Analysis*, 4, 224–234.

Ingersoll, J. E. (1987). *Theory of Financial Decision Making*, Rowman & Littlefield.

Izenman, A. J. (1991). "Recent Developments in Nonparametric Density Estimation," *Journal of the American Statistical Association*, 86, 205–224.

Jensen, M. C. (1968). "The Performance of Mutual Funds in the Period 1945–1964," *Journal of Finance*, 23, 389–416.

Kodde, D. A. and F. C. Palm (1986). "Wald Criteria for Jointly Testing Equality and Inequality Restrictions," *Econometrica*, 50, 1243–1248.

Kroll, D. and H. Levy (1980). "Stochastic Dominance Criteria: A Review and Some New Evidence," *Research in Finance*, Vol. 2, JAI Press, 163–227.

Kudo, A. (1963). "A Multivariate Analogue of the One-Sided-Test," *Biometrika*, 50, 403–418.

Kunsch, H. (1989). "The Jackknife and the Bootstrap for General Stationary Observations," *Annals of Statistics*, 17, 1217–1241.

Levy, H. (1992). "Stochastic Dominance and Expected Utility," *Management Science*, 38, 555–593.

Levy, H. and R. Brooks (1989). "An Empirical Analysis of Term Premiums Using Stochastic Dominance," *Journal of Banking and Finance*, 13, 245–260.

Levy, H. and G. Hanoch (1970). "Relative Effectiveness of Efficiency Criteria for Portfolio Selection," *Journal of Financial and Quantitative Analysis*, 5, 63–76.

Levy, H. and Y. Kroll (1978). "Ordering Uncertain Options with Borrowing and Lending," *Journal of Finance*, 33, 553–573.

Levy, H. and Y. Kroll (1979). "Efficiency Analysis with Borrowing and Lending: Criteria and Their Effectiveness," *Review of Economics and Statistics*, 61, 25–31.

Liu, R. Y. and K. Singh (1992). "Moving Blocks Jackknife and Bootstrap Capture Weak Dependence," in R. LePage, and L. Billard (ed.) *Exploring the Limits of Bootstrap*, John Wiley and Sons, 225–248.

Lo, Andrew W. (1983). "Long-Term Memory in Stock Market Prices," *Econometrica*, Vol. 59, 1279–1313.

Lo, A. W. and A. C. MacKinlay (1988). "Stock Market Prices Do Not Follow Random Walks: Evidence From A Simple Specification Test," *Review of Financial Studies*, 1, 41–46.

McFadden, D. (1989). "Testing for Stochastic Dominance," In T. B. Fomby and T. K. Seo (Ed.), *Studies in the Economics of Uncertainty: In Honor of Josef Hadar*, Springer Verlag, 113–134.

Mann, H. B. and A. Wald (1943). "On Stochastic Limit and Order Relationships," *Annals of Mathematical Statistics*, 14, 217–226.

Miller, R. G. (1981). *Simultaneous Statistical Inference*, 2nd ed., Wiley.

Mosteller, F. (1946). "On Some Useful 'Inefficient' Statistics," *Annals of Mathematical Statistics*, 17, 377–408.

Noreen, E. W. (1989). *Computer Intensive Methods for Testing Hypotheses*, John Wiley and Sons.

Parzen, E. (1962). "On Estimation of a Probability Density Function and Mode," *The Annals of Mathematical Statistics*, 33, 1065–1076.

Porter, R. (1978). "Portfolio Applications: Empirical Studies," in G. Whitmore and M. Findlay (eds.) *Stochastic Dominance*, Lexington Books, 117–161.

Porter, R. B., J. R. Wart and D. L. Ferguson (1973). "Efficient Algorithms for Conducting Stochastic Dominance Tests on Large Numbers of Portfolios," *Journal of Financial and Quantitative Analysis*, 8, 71–82.

Rao, C. R. (1965). *Linear Statistical Inference and Its Applications*, Wiley.

Reiss, R. D. (1981). "Nonparametric Estimation of Smooth Distribution Function," *Scandinavian Journal of Statistics*, 8, 116–119.

Richmond, J. (1982). "A General Method for Constructing Simultaneous Confidence Intervals," *Journal of the American Statistical Association*, 77, 455–460.

Roll, R. (1970). *The Behavior of Interest Rates: The Application of the Efficient Market Model to U.S. Treasury Bills*, Basic Books.

Roll, R. (1971). "Investment Diversification and Bond Maturity," *Journal of Finance*, 26, 51–66.

Roll, R. and S. A. Ross (1980). "An Empirical Investigation of the Arbitrage Pricing Theory," *Journal of Finance*, 5, 1073–1103.

Rothschild, M. and J. E. Stiglitz (1970). "Increasing Risk: I. A Definition," *Journal of Economic Theory*, 2, 225–243.

Savin, N. E. (1980). "The Bonferroni and Sheffe Multiple Comparison Procedures," *Review of Economic Studies*, 47, 255–273.

Savin, N. E. (1984). "Multiple Hypothesis Testing," In Z. Griliches and M. Intrilligator (eds.), *Handbook of Econometrics*, Vol. II, North-Holland.

Sen, P. K. (1972). "On the Bahadur Representation of Sample Quantiles for Sequences of ϕ-Mixing Random Variables," *Journal of Multivariate Analysis*, 2, 77–95.

Sharpe, W. F. (1966). "Mutual Fund Performance," *Journal of Business*, 39, 119–138.

Sheather, S. J. and J. S. Marron (1990). "Kernel Quantile Estimators," *Journal of the American Statistical Association*, 85, 410–416.

Shiller, R. (1990). "The Term Structure of Interest Rates," in B. M. Friedman and F. H. Hahn (ed.) *Handbook of Monetary Economics*, Vol. I, Chapter 13, 627–722.

Shorrocks, A. F. (1983). "Ranking Income Distributions," *Economica*, 50, 3-17.

Siddiqui, M. M. (1964). "Distribution of Quantiles in Samples from a Bivariate Population," *Journal of Research of the National bureau of Standards*, B. Mathematics and Mathematical Physics, Vol. 64B, No.3, 145-150.

Silverman, B. W. (1986). *Density Estimation for Statistics and Data Analysis*, Chapman and Hall.

Startz, R. (1982). "Do forecast Errors or Term Premia Really Make the Difference between Long and Short Rates?" *Journal of Financial Economics*, 10, 323–329.

Stein, W. E. and R. C. Pfaffenberger (1986). "Estimation of Error Probabilities in Stochastic Dominance," *Journal of Statistical Computation and Simulation*, 25, 137–165.

Stoline, M. R., and H. K. Ury (1979). "Table of the Studentized Maximum Modulus Distribution and an Application to Multiple Comparisons among Means," *Technometrics*, 21, 87–93.

Tapia, R. A., and J. R. Thompson (1978). *Nonparametric Probability Density Estimation*, Johns Hopkins University Press.

Tolley, H. D. and R. D. Pope (1988). "Testing for Stochastic Dominance," *American Journal of Agricultural Economics*, 70, 694–700.

Treynor, J. L. (1965). "How to Rate Management of Mutual Funds," *Harvard Business Review*, 43, 63–75.

Ullah, A. (1988). "Nonparametric Estimation of Econometric Functionals," *Canadian Journal of Economics*, 21, 625–658.

Vickson, R. G. (1977). "Stochastic Dominance Tests for Decreasing Absolute Risk-Aversion: II. General-Random Variables," *Management Science*, 23, 478–489.

Vickson, R. G. and Altman (1977). "On the Relative Effectiveness of Stochastic Dominance Rules: Extension to Decreasingly Risk-Averse Utility Functions," *Journal of Financial and Quantitative Analysis*, 12, 73–84.

Weiss, L. (1964). "On the Asymptotic Joint Normality of Quantiles From a Multivariate Distribution," *Journal of Research of the National bureau of Standards*, B. Mathematics and Mathematical Physics, Vol. 68B, No.2, 65–66.

Whitmore, G. A. (1970). "Third Degree Stochastic Dominance," *American Economic Review*, 60, 457–459.

Whitmore, G. A. and M. C. Findlay (Eds.) (1978). *Stochastic Dominance: An Approach to Decision-Making Under Risk*, Lexington Books.

Wilks, S. S. (1962). *Mathematical Statistics*, Wiley.

Wolak, F. A. (1989a). "Local and Global Testing of Linear and Nonlinear Inequality Constraints in Nonlinear Econometric Models," *Econometric Theory*, 5, 1–35.

Wolak, F. A. (1989b). "Testing Inequality Constraints in Linear Econometric Models," *Journal of Econometrics*, 41, 205–235.

Yang, Shie-Shien (1985). "A Smooth Nonparametric Estimator of a Quantile Function," *Journal of the American Statistical Association*, 80, 1004-1011.

CURRICULUM VITAE

Kuan **Xu**

EDUCATION:

1988- Concordia University, Montreal, Canada
—**Ph.D. Candidate** in Economics
—**Grade Point Average: 3.82**
—Specialization: Financial Economics, Econometrics, and Macroeconomics
—Concordia University and Departmental Scholarships
—Thesis Topic: Dominance Testing in Economics and Finance

1983-86 Concordia University, Montreal, Canada
—**M.B.A. Degree** in Business and Administration
—**Grade Point Average: 3.68**
—Specialization: Decision Science and International Trade

1976-78 Beijing Teachers' University, Beijing, China
—**Diploma** in Economics

HONORS:

—1988-94, University Scholarships and Teaching Fellowships, Concordia University
—1983-86, CIDA Scholarship, Concordia University
—1986, Dean's Honour List, Concordia University

WORK EXPERIENCE:

1988-1994 Concordia University
Department of Economics (1988-1994)
—**Lecturer:** Taught Statistics for Economists I and II, Mathematics for Economists I and II, Statistical Packages and Economic Data, and Econometrics I and II.
—**Research Assistant:** Participated in the research on the performance of the Canadian equity portfolios and the asset pricing models.
—**Research Assistant:** Programmed in FORTRAN for model simulation and GAUSS for nonparametric estimation.
—**Teaching Assistant:** Tutored graduate students TSP and SHAZAM.

and undergraduate students statistics and mathematics.

1986–88 People's University, Beijing, China
Department of Economics
—**Lecturer:** Taught Macroeconomics, Economic Development and Econometrics.
—**Researcher:** Conducted the short-term forecasting with the People's Bank of China, the Bank of China, and other financial institutions.

1978–83 People's University, Beijing, China
Department of Economics
—**Instructor:** Taught Introductory Economics.
—**Researcher:** Conducted research on China's economic reform issues.

COMPUTER SKILLS:

1. Language: FORTRAN, GAUSS
2. System: VAX/VMS, DOS, WINDOWS
3. Econometrics/Statistical Software: SAS, TSP, SHAZAM, MINITAB
4. Spread Sheet: LOTUS, EXCEL
5. Word Processor: LATEX, CHIWRITER, WORD, WORD PERFECT, SCIENTIFIC WORD

PUBLICATIONS:

1. **Books:**
   - (a) (with Song Ligang) *Foreign Exchange: Theories and Forecasting*, The People's University Press, 1990, China.
   - (b) (with Wu Qi, Wang Songbao, and Tao Wenda) *Industrialization and Growth: A Comparative Study* by H. Chenery, S. Robinson, and M. Syrquin, Chinese Edition, Shanghai San Lian Press, 1989, China.
2. **Papers:**
   - (a) (with Liu Yipeng) "Growth, Fluctuations and Mechanism—Exploring the Operational Mechanism of China's Economy", *Economic Theory and Business Management*. No.6, 1987, China.

(b) (with V. H. Kirpalani) "Effective International Market Potential Assessment: China", in Melanie Wallendorf and Paul Anderson (ed.), *Advances in Consumer Research*, Volume XIV, 398-402, The American Association of Consumer Research, 1987.

CONFERENCES:

1. The paper, "New Distribution-Free Tests for Stochastic Dominance," was presented at the Annual Conference of the Canadian Economic Association, Calgary, Canada, 1994.
2. The Canadian Macroeconomics Study Group Meeting, Victoria, Canada, 1993.
3. The paper, "Implicit and Forward Tax Rate Determination Using Yield-Spread Equilibrium Models for Taxable and Tax-Exempt Bonds," was presented at the Northern Finance Association Conference, Montreal, Canada, 1991.
4. International Monetary Fund and the People's Bank of China 1986 Macroeconomic Management Symposium, Beijing, China, 1986.
5. The paper, "Effective International Market Potential Assessment: China," was presented at the Annual Conference of the American Association of Consumer Research, Toronto, Canada, 1986.

RESEARCH GRANTS:

1. Concordia University Professional Development Fund, 1992-1993.
2. People's University Financial Research Fund, 1987-1988.

PROFESSIONAL ASSOCIATIONS:

1. Member, Canadian Economic Association.
2. Member, American Economic Association.
3. Member, Financial Management Association.