Stochastic Dominance via Quantile Regression[*]

Working Paper Series – 11-01

Pin Ng

Franke College of Business

Northern Arizona University

Flagstaff, AZ 86011-5066

Wing-Keung Wong

Department of Economics

Hong Kong Baptist University

WLB, Shaw Campus, Kowloon Tong

Hong Kong

Zhijie Xiao

Department of Economics

Boston College

Chestnut Hill, MA 02467 USA

Abstract

We derive a new way to test for stochastic dominance between the return of two assets using a quantile regression formulation. The test statistic is a variant of the one-sided Kolmogorov-Smirnoff statistic and has a limiting distribution of the standard Brownian bridge. We also illustrate how the test statistic can be extended to test for stochastic dominance among k assets. This is useful when comparing the performance of individual assets in a portfolio against some market index. We show how the test statistic can be modified to test for stochastic dominance up to the α-quantile in situation where the return of one asset does not dominate another over the whole spectrum of the return distribution.

Keywords: Quantile regression, stochastic dominance, Brownian bridge, test statistic.

1   Introduction

Stochastic dominance finds applications in many areas. In finance, it is used to assess portfolio diversification, capital structure, bankruptcy risk, and option's price bound. In welfare economics, it is used to measure income distribution and income inequality In reinsurance coverage, the insured use it to select the best coverage option while the insurers use it to assess whether the options are consistently priced. It is also used to select effective treatment in medicine and selection of the best irrigation system in agriculture.

There are two big classes of stochastic dominance tests. The first is based on the inf / sup statistics over the support of the distributions as in McFadden (1989), Klecan, McFadden and McFadden (1991), and Kaur, Rao and Singh (1994). The second class is based on comparison of the distributions over a set of grid points as in Anderson (1996), Dardanoni and Forcina (1998, 1999), and Davidson and Duclos (2000). We derive a new way to test for stochastic dominance using a quantile regression formulation. The test

statistic is a variant of the one-sided Kolmogorov-Smirnoff statistic and has a limiting distribution of the standard Brownian bridge. It is based on the inf / sup statistics over the range of the distribution. So, our test can be considered a variant of the first class that is based on the inf / sup statistics over the support of the distributions. We illustrate how our test can be performed by casting the problem in a quantile regression setting. As a result, our test statistic is user-friendly because it can be computed by adapting existing statistical software that performs quantile regression estimation, e.g., the quantreg package in R, the PROC QUANTREG in SAS/STAT, the qreg in STATA, the LAD procedure in TSP and the quantile regression model in LIMDEP.

2 Testing for Stochastic Dominance via Quantile Regression

We consider testing stochastic dominance between two return distributions in Section 2.1. We will test stochastic dominance up to the q-th quantile in Section 2.2. Extension to k return distributions is provided in Section 2.3.

2.1 Stochastic Dominance Between Two Unconditional Distributions

We want to compare two populations on the basis of two samples $\{y_{11}, \cdots, y_{1n_1}\}$ and $\{y_{21}, \cdots, y_{2n_2}\}$, with distribution functions F_1 and F_2, respectively. Here, $y_{11}, \cdots, y_{1n_1}$ might be return observations from the first asset, i.e. realizations of R_1, and $y_{21}, \cdots, y_{2n_2}$ might be return observations from the second asset, i.e. realizations of R_2.

Suppose that the two random samples $\{y_{1t}\}_{t=1}^{n_1}$ and $\{y_{2t}\}_{t=1}^{n_2}$ have location parameters

3

α_1 and α_2, respectively:

$$\begin{cases} y_{1t} = \alpha_1 + u_t, \ t = 1, ..., n_1, \\ y_{2t} = \alpha_2 + v_t, \ t = n_1 + 1, ..., n, \end{cases} \tag{1}$$

where u_t and v_t are mean zero with distribution functions F_u and F_v, respectively. The distribution function (CDF) of y_{1t} is given by

$$F_1(y) = \Pr\left(\alpha_1 + u_t < y\right) = F_u\left(y - \alpha_1\right)$$

and the distribution function (CDF) of y_{2t} is given by

$$F_2(y) = \Pr\left(\alpha_2 + v_t < y\right) = F_v\left(y - \alpha_2\right).$$

R_1 is said to stochastically dominate R_2 at first order, denoted by $R_1 \succ_1 R_2$, if

$$F_1(y) \leq F_2(y) \text{ for all } y \text{ and there exists } y \text{ such that } F_1(y) < F_2(y). \tag{2}$$

This is equivalent to

$$Q_1(\tau) = F_1^{-1}(\tau) \geq F_2^{-1}(\tau) = Q_2(\tau), \text{ for all } \tau \in (0,1)$$

with strict inequality on at least one point.

To test stochastic dominance, we introduce a dummy variable D_t defined as follows:

$$D_t = \begin{cases} 1, \text{ for } t = 1, ..., n_1, \\ 0, \text{ for } t = n_1 + 1, ..., n \end{cases}. \tag{3}$$

4

Thereafter, we combine the return data from the two assets as follows:

$$y_t = \begin{cases} y_{1t}, & \text{for } t = 1, ..., n_1, \\ y_{2t}, & \text{for } t = n_1 + 1, ..., n \end{cases} \qquad (4)$$

Then, we obtain the following pooled model:

$$y_t = \alpha + \beta D_t + w_t = z_t'\theta + w_t, \qquad (5)$$

where $z_t = (1, D_t)'$, $\theta = (\alpha, \beta)'$, $w = v_t + (u_t - v_t)D_t$, $\alpha = \alpha_2$, and $\beta = \alpha_1 - \alpha_2$.

The quantile function of the returns can be written as

$$Q_{y_t}(\tau|D_t) = \alpha(\tau) + \beta(\tau)D_t = z_t'\theta(\tau) = Q_{y_t}(\tau|z_t), \qquad (6)$$

where $\alpha(\tau) = \alpha + Q_v(\tau)$, $\beta(\tau) = \beta + Q_u(\tau) - Q_v(\tau)$, $\theta(\tau) = (\alpha(\tau), \beta(\tau))'$, and the conditional distribution function of y_t is $P(y_t < y|z_t) = F_{y_t}(y|z_t) = F_t(y)$. Notice that

$$Q_{y_t}(\tau|D_t = 1) = \alpha_1 + Q_u(\tau) = Q_1(\tau) = F_{y_t}^{-1}\left(\tau|z_t = (1,1)'\right),$$

and

$$Q_{y_t}(\tau|D_t = 0) = \alpha_2 + Q_v(\tau) = Q_2(\tau) = F_{y_t}^{-1}\left(\tau|z_t = (1,0)'\right).$$

The hypothesis that R_1 stochastically dominates R_2 at first order can then be expressed as

$$Q_{y_t}(\tau|D_t = 1) \geq Q_{y_t}(\tau|D_t = 0) \text{ for all } \tau$$

with strict inequality on at least one point.

It is easy to see that the regression quantile process, $\beta(\tau)$, of the dummy variable D_t,

measures the distributional difference between the two groups (assets) in that

$$\beta(\tau) = [\alpha_1 + Q_u(\tau)] - [\alpha_2 + Q_v(\tau)] = Q_{y_t}(\tau|D_t = 1) - Q_{y_t}(\tau|D_t = 0) = Q_1(\tau) - Q_2(\tau).$$

Hence, the first order stochastic dominance can be further re-formulated as, in the pooled quantile regression model (6),

$$\beta(\tau) \geq 0, \text{ for all } \tau$$

with strict inequality on at least one point.

As a result, we may construct statistical tests for stochastic dominance based on the following quantile regression:

$$\min_{\alpha,\beta} \sum_{t=1}^{n} \rho_\tau(y_t - \alpha - \beta D_t) \tag{7}$$

where $\rho_\tau(u) = u\big(\tau - I(u < 0)\big)$ is the "check function" as in Koenker and Bassett (1978).

We consider the null hypothesis that there is no difference between the distributions of the two assets, against the alternative of first order dominance. Thus, our null hypothesis, $H_0 : F_1(y) = F_2(y)$, for all y; $\big($or $Q_1(\tau) = Q_2(\tau)$, for all $\tau \in (0,1)\big)$ can be expressed as:

$$H_0 : \beta(\tau) = 0, \text{ for all } \tau \in (0,1). \tag{8}$$

For the alternatives of first order dominance, $R_1 \succ_1 R_2$, we can consider:

$$H_{1A} : R_1 \succ_1 R_2 ,$$
$$F_1(y) \leq F_2(y) \text{ for all } y,$$
$$\text{or } Q_1(\tau) \geq Q_2(\tau) \text{ for all } \tau \in (0,1),$$
$$\text{or } \beta(\tau) \geq 0 \text{ for all } \tau \in (0,1)$$

6

with strict inequality on at least one point.

Similarly, the alternatives of first order dominance, $R_2 \succ_1 R_1$, can be expressed as :

$$H_{1B} : R_2 \succ_1 R_1 \ ,$$

$$F_1(y) \geq F_2(y) \text{ for all } y,$$

$$\text{or } Q_1(\tau) \leq Q_2(\tau) \text{ for all } \tau \in (0, 1),$$

$$\text{or } \beta(\tau) \leq 0 \text{ for all } \tau \in (0, 1)$$

with strict inequality on at least one point.

We construct our tests based on the following one-sided Kolmogorov-Smirnoff statistics

$$\inf_{\tau} \widehat{\beta}(\tau) \text{ for } H_{1A}, \tag{9}$$

and

$$\sup_{\tau} \widehat{\beta}(\tau) \text{ for } H_{1B}, \tag{10}$$

where $\widehat{\beta}(\tau)$ is the regression quantile process obtained from performing (7).

Under the null, $\widehat{\beta}(\tau)$ should be close to 0, but under the alternative hypothesis H_{1A} that $R_1 \succ_1 R_2$, $\inf_{\tau} \beta(\tau)$ should be greater than zero ($\inf_{\tau} \beta(\tau) > 0$). Thus, the null hypothesis should be rejected in favor of the hypothesis that $R_1 \succ_1 R_2$ if $\inf_{\tau} \widehat{\beta}(\tau)$ is significantly positive.

Similarly, under the alternative hypothesis H_{1B} that $R_2 \succ_1 R_1$, $\sup_{\tau} \beta(\tau)$ should be less than zero ($\sup_{\tau} \beta(\tau) < 0$). Thus, the null hypothesis should be rejected in favor of the alternative hypothesis that $R_2 \succ_1 R_1$ if $\sup \widehat{\beta}(\tau)$ is significantly negative.

For asymptotic analysis, we assume that there exists $0 < \delta_1 < \delta_2 < 1$ so that $0 < \delta_1 < n_i/n < \delta_2 < 1$, for $i = 1, 2$, where $n = n_1 + n_2$. To study the asymptotic properties of the testing procedures, it is convenient for us to make the following assumptions:

Assumption 1: $\{y_t\}$ are uncorrelated across t and the conditional distribution function

7

F_t has a continuous Lesbesgue density, f_t, with $f_t(u) > 0$ on $\{u : 0 < F_t(u) < 1\}$.

Assumption 2: As $n \to \infty$, $\Omega_n = n^{-1} Z_n' Z_n \to \Omega_0$, where Ω_0 is a positive definite matrix.

The limiting distribution of $\widehat{\beta}(\tau)$, however, contains nuisance parameters, and thus, the statistics in (9) or (10) cannot be directly used as testing statistics for stochastic dominance. In particular, the limiting regression quantile process is a Gaussian process whose covariance matrix depends on the sparsity function $\varphi(\tau) \equiv f(F^{-1}(\tau))$ and the limiting matrix Ω_0. In order to obtain asymptotically-distributional-free tests, we need to estimate the matrix Ω_0 and the function $\varphi(\tau)$, and re-standardize the regression quantile process using these estimates. The matrix Ω_0 can be naturally estimated by $\Omega_n = n^{-1} Z_n' Z_n$, and the sparsity function can be estimated using existing methods in the literature.

We impose the following assumption for the sparsity estimator. For more study on the sparsity estimator, see, e.g. Portnoy and Koenker (1989).

Assumption 3: $\widehat{\varphi}(s)$ is a uniformly consistent estimator of $\varphi(\cdot)$ over $\tau \in \Pi$, where $\Pi = [\epsilon, 1 - \epsilon]$, and ϵ is a small positive number.

We define
$$\widehat{V}_1(\tau) = \sqrt{n} \widehat{f}(F^{-1}(\tau)) \left[R \Omega_n^{-1} R' \right]^{-1/2} \widehat{\beta}(\tau),$$
where $R = (0, 1)$ is a row vector, and $\widehat{f}(F^{-1}(\tau))$ is a consistent estimator of $f(F^{-1}(\tau))$. We summarize the asymptotic results in the following theorem:

Theorem 1 *Under the null hypothesis in (8) and under Assumptions 1 to 3,*

$$\sup_{\tau \in \Pi} \widehat{V}_1(\tau) \Rightarrow \sup_{\tau \in \Pi} \overline{W}_1(\tau), \text{ and } \inf_{\tau \in \Pi} \widehat{V}_1(\tau) \Rightarrow \inf_{\tau \in \Pi} \overline{W}_1(\tau)$$

where $\overline{W}_1(\tau)$ is a one-dimensional Brownian bridge.

We note that the limiting distribution of the above one-sided Kolmogorov-Smirnov statistic is free of nuisance parameters and thus it can be used in testing for stochastic dominance.

Remark 1   *Like the construction of many other Kolmogorov type tests, we consider superium of the process over $\Pi = [\epsilon, 1 - \epsilon]$. This is because the estimation of the sparsity function becomes poor as $\tau \to 0$ or 1. Thus, $\widehat{f}(F^{-1}(\tau))$ is a uniformly consistent estimator of $f(F^{-1}(\tau))$ uniformly over $[\epsilon, 1 - \epsilon]$ but the uniform convergence can hardly hold on $[0, 1]$, see our discussion on related issues in Section 2.4 in which higher order stochastic dominance is investigated. In this situation, a weaker condition can be used and it is possible to consider superium of the process over $[0, 1]$.*

2.2   Stochastic Dominance Up to the q-th Quantile

In situation where the return of one asset does not dominate another over the entire spectrum of the return distribution, we define stochastic dominance up to the q-th quantile, denoted by $R_1 \succ_1^q R_2$, if

$$Q_1(\tau) = F_1^{-1}(\tau) \geq F_2^{-1}(\tau) = Q_2(\tau), \text{ for all } \tau \leq q \tag{11}$$

with strict inequality on at least one point.

The proposed concept of stochastic dominance up to a specified quantile is a useful concept because we might be particularly interested in stochastic dominance over a certain range (say, left tail) of the distribution in some applications. Such a concept is parallel to the stochastic dominance up to a poverty line z studied by Davidson and Duclos (2000). In particular, if we define the dominance of R_1 over R_2 up to level z, denoted by $R_1 \succ_1^z R_2$, when $F_1(x) \leq F_2(x)$ for all $x < z$ with strict inequality for at least one point, then $R_1 \succ_1^q R_2$ implies $R_1 \succ_1^z R_2$ for $z = \max\left(F_1^{-1}(z), F_2^{-1}(z)\right)$ and $R_1 \succ_1^z R_2$ implies $R_1 \succ_1^q R_2$ for $q = max\left(F_1(z), F_2(z)\right)$.

We can test stochastic dominance up to the q-th quantile based on

$$\sup_{\tau \leq q} \widehat{V}_1(\tau), \text{ and } \inf_{\tau \leq q} \widehat{V}_1(\tau),$$

which has limiting distribution:

$$\sup_{\tau \leq q} \widehat{V}_1(\tau) \Rightarrow \sup_{\tau \leq q} \overline{W}_1(\tau), \text{ and } \inf_{\tau \leq q} \widehat{V}_1(\tau) \Rightarrow \inf_{\tau \leq q} \overline{W}_1(\tau).$$

2.3   Generalizations to k Assets

One may be interested in evaluating the performance of k assets or to evaluate whether one asset or portfolio, say for example, the market index, outperforms the remaining $(k-1)$ assets. In this situation, one may be interested in testing the dominance relationship among the k assets and find the partial order (see, for example, Egozcue and Wong, 2010) among them.

As in the two-asset case, we can combine these assets by using dummy variables. Suppose that there are return observations of k assets $\{y_{jt}\}_{t=1}^{n_j}$, $j = 1, ..., k$,

$$y_{jt} = \alpha_j + u_{jt}, \ t = 1, ..., n_j, \ j = 1, ..., k, \tag{12}$$

where u_{jt} are zero mean with distribution function $F_{uj}(\cdot)$ and quantile function $Q_{uj}(\cdot)$. In addition, we denote the CDF and quantile functions for the return distribution of the j-th asset as $F_j(\cdot)$ and $Q_j(\cdot)$.

To investigate the partial ordering of return distributions, we define the following $k-1$ dummy variables

$$D_{jt} = \begin{cases} 1, \text{ for the } j\text{-th subsample}, \\ 0, \text{ otherwise}, \end{cases}, \ j = 1, ..., k-1. \tag{13}$$

We can pool the k assets together and consider the following regression model

$$y_t = z_t'\theta + w_t, \ t = 1, ..., n,$$

where

$$z_t = (D_{1t}, ..., D_{k-1,t}, 1), \text{ and } \theta = (\theta_1, ..., \theta_k)' = (\alpha_1 - \alpha_k, \cdots, \alpha_{k-1} - \alpha_k, \alpha_k)'.$$

Again, we define

$$Q_{y_t}(\tau | D_{jt} = 1) = \alpha_j + Q_{uj}(\tau) = Q_j(\tau), \ j = 1, ..., k-1$$

and

$$Q_{y_t}(\tau | D_{jt} = 0, \ j = 1, ..., k-1) = \alpha_k + Q_{uk}(\tau) = Q_k(\tau).$$

Thus, the hypothesis that the return distribution of R_j $(j = 1, ..., k-1)$ dominates that of R_k at first order can be expressed as

$$Q_{y_t}(\tau | D_{jt} = 1) \geq Q_{y_t}(\tau | D_{jt} = 0, \ j = 1, ..., k-1), \text{ for all } \tau.$$

Again, since the regression quantile process of the dummy variable D_{jt} measures the distributional difference between the asset j and asset k,

$$\theta_j(\tau) = [\alpha_j + Q_{uj}(\tau)] - [\alpha_k + Q_{uk}(\tau)] = Q_{y_t}(\tau | D_{jt} = 1) - Q_{y_t}(\tau | D_{jt} = 0, \ j = 1, ..., k-1),$$

the first order stochastic dominance of asset k by j can be further re-formulated as, in the pooled quantile regression model (6),

$$\theta_j(\tau) \geq 0, \text{ for all } \tau.$$

11

and we can test for stochastic dominance in a similar way.

The choice of asset k usually will depend on the application. We may consider choosing some index or other benchmark asset as asset k, and compare other stocks with this index.

2.4   Testing for Higher Order Stochastic Dominance

Second order stochastic dominance can be tested based on integrals of $\widehat{\beta}(\tau)$ or more precisely, integrals of $\widehat{V}(\tau)$.

From the definition of the first order stochastic dominance in (2), one could define the second order stochastic dominance, $R_1 \succ_2 R_2$ if

$$\int_{-\infty}^{y} F_1(t)dt \leq \int_{-\infty}^{y} F_2(t)dt \text{ for all } y \tag{14}$$

with strict inequality on at least one point. One could easily show that this definition is equivalent to the following:

$$\int_{0}^{\tau} Q_1(s)ds \geq \int_{0}^{\tau} Q_2(s)ds, \text{ for all } \tau \in (0,1) \tag{15}$$

with strict inequality on at least one point.

Theorem 2   *Let y_1 and y_2 be random variables with distribution functions F_1 and F_2 respectively and let $Q_1(\tau) = F_1^{-1}(\tau)$ and $Q_2(\tau) = F_2^{-1}(\tau)$ for each $\tau \in (0,1)$. Under Assumption 1,*

$$\int_{-\infty}^{y} F_1(t)dt \leq \int_{-\infty}^{y} F_2(t)dt \text{ for all } y$$

is equivalent to

$$\int_{0}^{\tau} Q_1(s)ds \geq \int_{0}^{\tau} Q_2(s)ds, \text{ for all } \tau \in (0,1).$$

Define

$$\widehat{V}_2(\tau) = \int_0^\tau \widehat{V}_1(s)ds = \sqrt{n}\left[R\Omega_0^{-1}R'\right]^{-1/2}\int_0^\tau \widehat{f}(F^{-1}(s))\widehat{\beta}(s)ds,$$

where $\widehat{\beta}(s)$ is the regression quantile process of (5), $\widehat{f}(F^{-1}(s))$ is an estimator of $f(F^{-1}(s))$. If we consider the null hypothesis that there is no difference between the distributions of the two assets, i.e. H_0: $Q_1(\tau) = Q_2(\tau)$, for all $\tau \in (0, 1)$, against the alternative of second order dominance, i.e. H_{2A}: $\int_0^\tau Q_1(s)ds \geq \int_0^\tau Q_2(s)ds$, for all $\tau \in (0, 1)$, we may construct a test based on the following one-sided Kolmogorov-Smirnoff statistic

$$\inf_\tau \widehat{V}_2(\tau).$$

Similarly, if we test the null against the alternative that the return distribution of R_2 dominates that of R_1 in second order, i.e. H_{2B}: $\int_0^\tau Q_1(s)ds \geq \int_0^\tau Q_2(s)ds$, for all $\tau \in (\mathbf{0}, \mathbf{1})$, we may construct a test based on

$$\sup_\tau \widehat{V}_2(\tau).$$

For testing higher order stochastic dominance, we only need the following weaker condition in place of Assumption 3.

Assumption 3′: Let $\hat{f}(F^{-1}(s))$ be an estimator of $f(F^{-1}(s))$ such that
$\int_0^\tau \left[\hat{f}(F^{-1}(s)) - f(F^{-1}(s))\right]\widehat{\beta}(s)ds = o_p(1)$ uniformly in $\tau \in [0, 1]$.

The condition: $\int_0^\tau \left[\hat{f}(F^{-1}(s)) - f(F^{-1}(s))\right]\widehat{\beta}(s)ds = o_p(1)$ uniformly in $\tau \in [0, 1]$ is much weaker than $\sup_{0 \leq \tau \leq 1}\left|\hat{f}(F^{-1}(s)) - f(F^{-1}(s))\right| = o_p(1)$. Under this assumption and Assumptions 1-2, we may construct a quantile regression based test for higher order stochastic dominance. For example, the following test for second order stochastic dominance can be constructed:

Theorem 3   *Under the null hypothesis in (8) and under Assumptions 1 to 3′,*

$$\sup_{\tau \in [0,1]} \widehat{V}_2(\tau) \Rightarrow \sup_{\tau \in [0,1]} \int_0^\tau \overline{W}_1(s)ds, \ and \ \inf_{\tau \in [0,1]} \widehat{V}_2(\tau) \Rightarrow \inf_{\tau \in [0,1]} \int_0^\tau \overline{W}_1(s)ds.$$

Similar analysis can be carried over to third order stochastic dominance.

Table 1: Critical values of $\inf_\tau \widehat{\beta}(\tau)$ for the upper-tail test on H_{1A}.

	α				
ϵ	0.001	0.005	0.010	0.050	0.100
0.05	0.22	0.11	0.05	-0.10	-0.20
0.10	0.36	0.23	0.16	-0.04	-0.15
0.15	0.51	0.34	0.26	0.03	-0.09
0.20	0.63	0.43	0.35	0.10	-0.03
0.25	0.73	0.54	0.44	0.18	0.03
0.30	0.87	0.65	0.55	0.26	0.10

3   Simulation of the Asymptotic Critical Values

The following procedure is used to obtain the α-level critical value of $\inf \widehat{V}_1$ in the upper-tail test of H_{1A}: (i) generate a one-dimensional Brownian bridge over a grid of N+1 points between 0 and 1, (ii) obtain the maximum value of the bridge over $\Pi = [\epsilon, 1 - \epsilon]$ for $\epsilon = 0.05$ to 0.30 in increment of 0.05, (iii) repeat step (i) and (ii) for 40,000 times, (iv) obtain the $(1 - \alpha)$-th quantile of the minimum values. Following DeLong (1981) and Andrews (1993), we use $N = 3600$ to approximate the one-dimensional Brownian bridge. The level of significance α is chosen to be $0.1\%, 0.5\%, 1\%, 5\%, 10\%$. The critical value of the lower-tail test is obtained similarly except the α-the quantile of the maximum values is used instead in step (iv). The critical values for the upper-tail test of H_{1A} are presented in Table 1 while Table 2 contains those for the lower-tail test of H_{1B}. We can see that the critical values of the lower-tail test are basically the negative images of the upper-tail test due to symmetry.

Figure 1 shows 500 simulated paths of the one-dimensional Brownian bridge while Figure 2 contains 500 simulated path of the $\widehat{V}_1(\tau)$ test statistic for first order stochastic dominance. We can see that the paths of $\widehat{V}_1(\tau)$ behave like a one-dimensional Brown bridge but the simulated 90% confidence band of $\widehat{V}_1(\tau)$ is slightly narrower than that of

Table 2: Critical values of $\sup_\tau \widehat{\beta}(\tau)$ for the lower-tail test on H_{1B}.

ϵ	α				
	0.001	0.005	0.010	0.050	0.100
0.05	-0.21	-0.11	-0.05	0.10	0.20
0.10	-0.37	-0.23	-0.15	0.04	0.15
0.15	-0.48	-0.33	-0.25	-0.03	0.09
0.20	-0.59	-0.43	-0.34	-0.10	0.03
0.25	-0.76	-0.53	-0.44	-0.17	-0.03
0.30	-0.87	-0.65	-0.54	-0.25	-0.10

the Brownian bridge and the simulated variance of $\hat{V}_1(\tau)$ is also slightly smaller than the simulated variance of the Brownian bridge.

4   Some Simulation Results on the Empirical Level of Significance and Power

To study the empircal size of $\hat{V}_1(\tau)$, we conducted two simulations where F_1 and $F_2 \sim N(0,1)$, and F_1 and $F_2 \sim$Burr(4.7,0.55). The *Burr* distribution is commonly used in stochastic dominance simulation studies as in Dardanoni and Forcina (1999), and Tse and Zhang (2004). The nominal level we used in the studies is 5% and the number of replications is 1000. The results of the empirical levels are presented in Table 3 and Table 4. In general, the larger the sample size, the more trimming is needed.

To study the power of the test, we perform the following simulations:

Model 1: $F_1 \sim N(1,1)$ and$F_2 \sim N(-1,1)$

Model 2: $F_1 \sim Burr(4.7, 0.55)$ and$F_2 \sim Burr(4.7, 0.95)$

Model 3: $F_1 \sim Burr(4.7, 0.55)$ and$F_2 \sim Burr(-4.7, 0.65)$

The power of the test decreases as we move from Model 1 to Model 3. This is due to the decreasing separation between the two CDFs when we move from Model 1 to 3 as

16

Brownian Bridge: N = 100, Nmc = 500
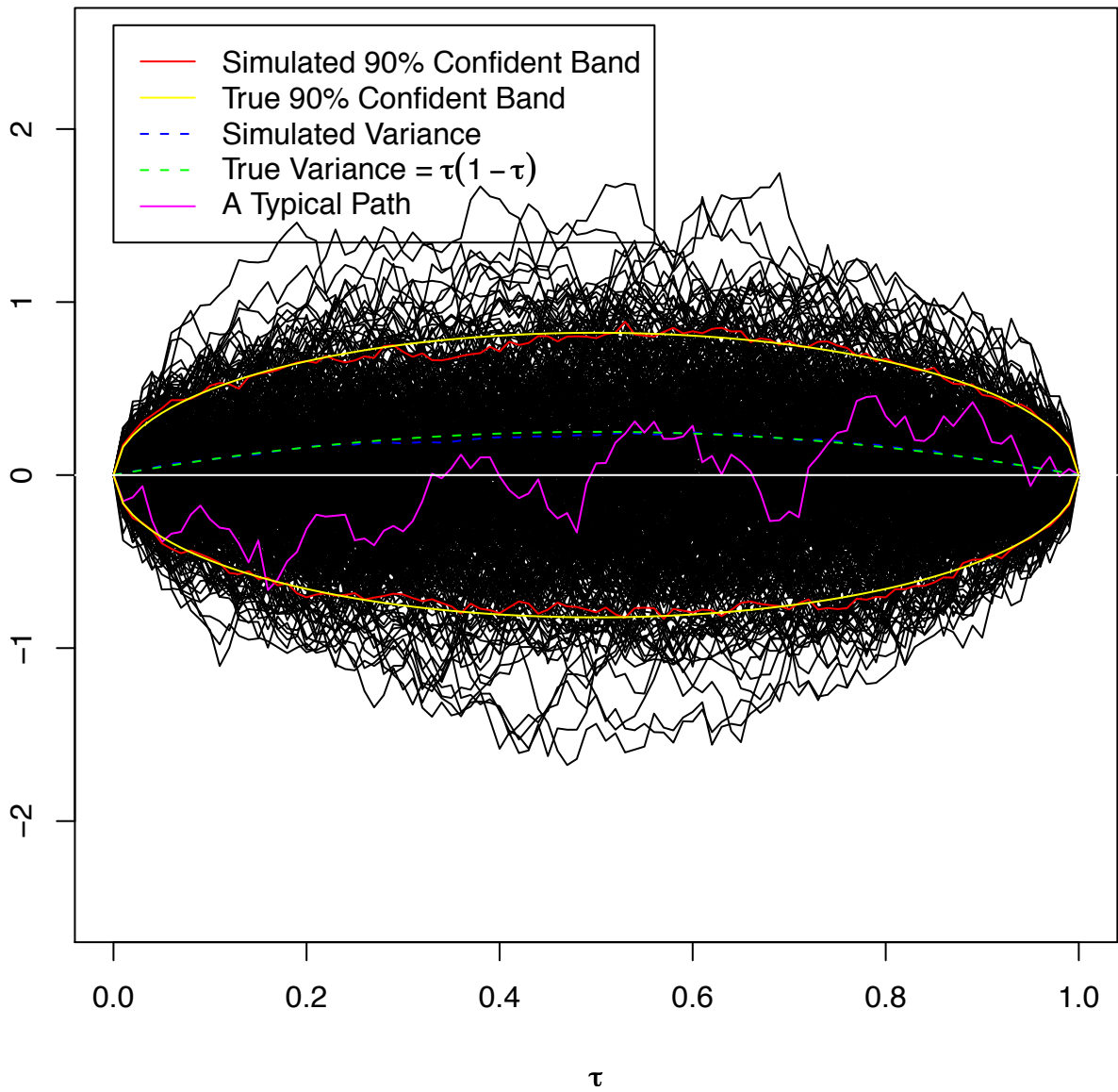
Figure 2: Simulated Paths of $\hat{V}_1(\tau)$

$\hat{V}_1(\tau)$: N = 100, Nmc = 500

Legend:
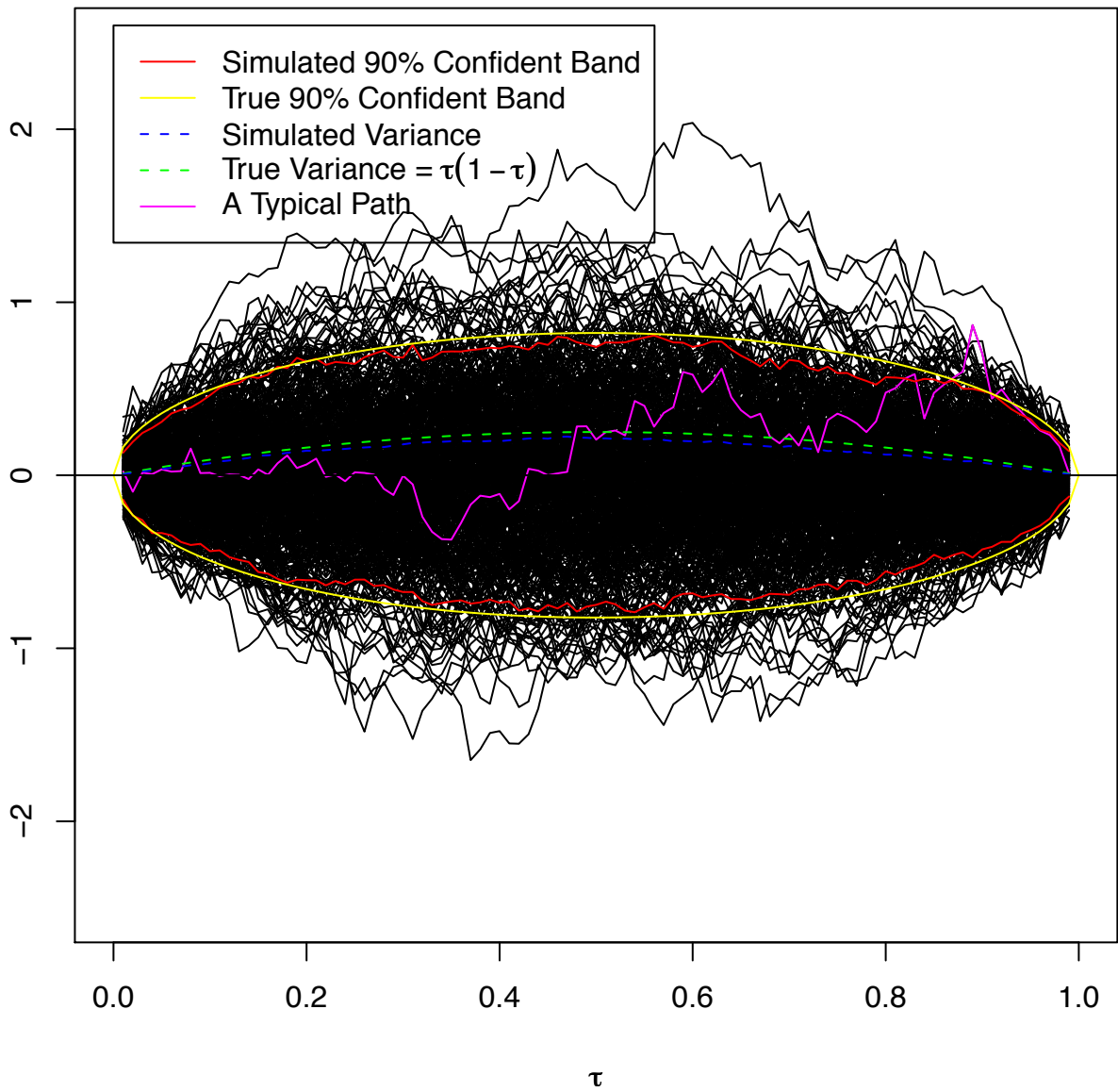- Simulated 90% Confident Band
- True 90% Confident Band
- Simulated Variance
- True Variance = $\tau(1-\tau)$
- A Typical Path

Table 3: Empirical size of $\inf\limits_{\tau\in\Pi}\hat{V}_1(\tau)$ for $F_1$and$F_2 \sim N(0,1)$

	n			
ϵ	50	100	250	500
.05	.077	.092	.082	.076
.10	.069	.069	.083	.065
.15	.062	.043	.064	.056
.20	.048	.043	.052	.050
.25	.056	.042	.057	.040
.30	.046	.043	.05	.046

Table 4: Empirical size of $\inf\limits_{\tau\in\Pi}\hat{V}_1(\tau)$ for $F_1$and$F_2 \sim Burr(4.7, 0.55)$

	n			
ϵ	50	100	250	500
.05	.075	.088	.097	.079
.10	.081	.101	.062	.067
.15	.048	.059	.063	.055
.20	.052	.058	.056	.054
.25	.049	.044	.049	.054
.30	.048	.039	.058	.063

Table 5: Power of $\inf_{\tau \in \Pi} \hat{V}_1(\tau)$ for Model 1.

	n			
ϵ	50	100	250	500
.05	1.00	1.00	1.00	1.00
.10	1.00	1.00	1.00	1.00
.15	1.00	1.00	1.00	1.00
.20	1.00	1.00	1.00	1.00
.25	1.00	1.00	1.00	1.00
.30	1.00	1.00	1.00	1.00

Table 6: Power of $\inf_{\tau \in \Pi} \hat{V}_1(\tau)$ for Model 2.

	n			
ϵ	50	100	250	500
.05	0.638	0.828	0.934	0.977
.10	0.640	0.839	0.956	0.995
.15	0.575	0.813	0.975	0.998
.20	0.596	0.806	0.978	1.000
.25	0.537	0.834	0.991	0.999
.30	0.593	0.831	0.990	1.000

is evident from Figure 3, 4, and 5. In general, the power increases as the sample size increases.

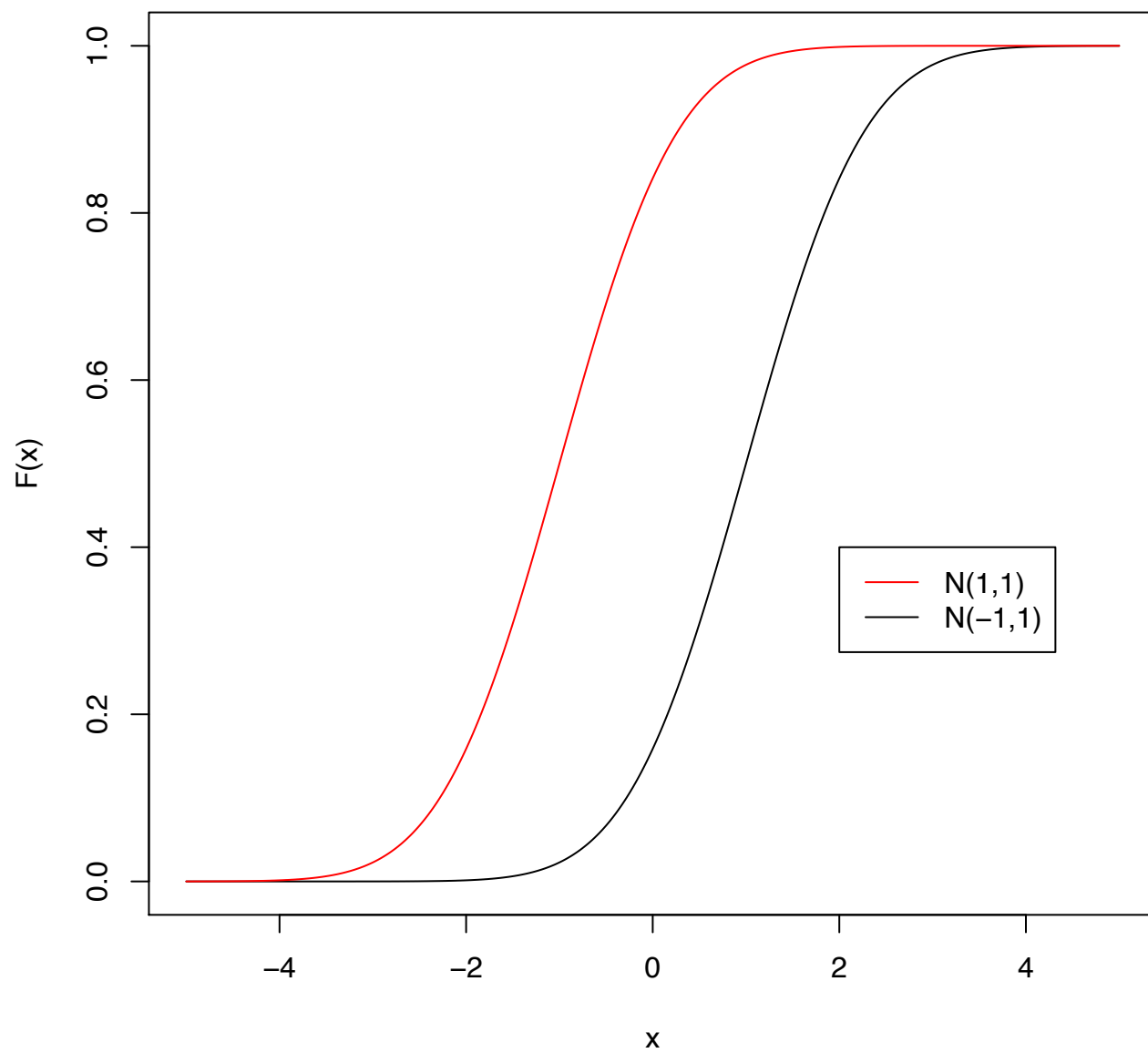Figure 3: Cumulative distribution functions of Model 1.

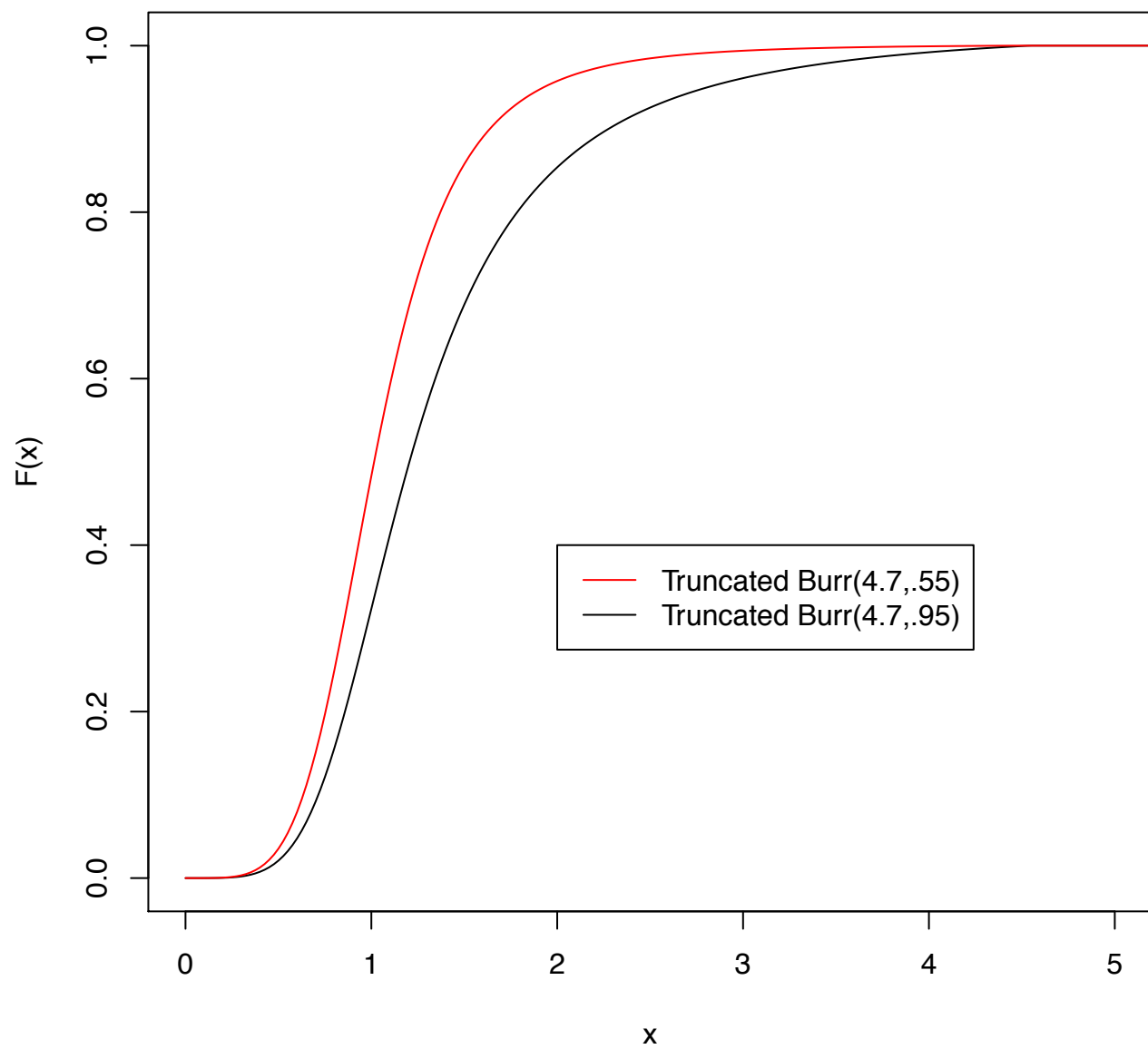Figure 4: Cumulative distribution functions of Model 2.
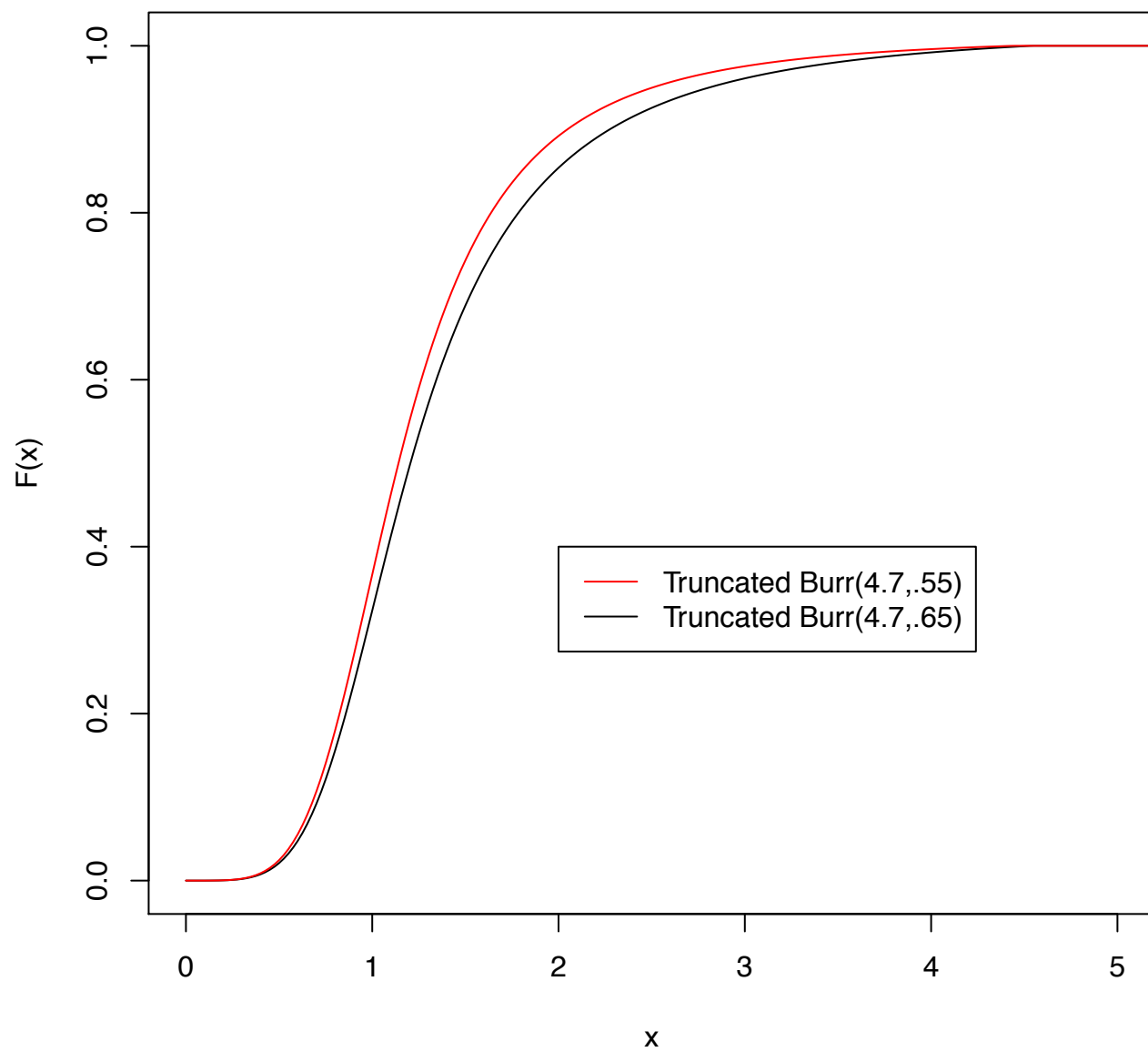
Figure 5: Cumulative distributions of Model 3.

Truncated Burr(4.7,.55)
Truncated Burr(4.7,.65)

Table 7: Power of $\inf_{\tau \in \Pi} \hat{V}_1(\tau)$ for Model 3.

			n	
ϵ	50	100	250	500
.05	0.196	0.308	0.410	0.592
.10	0.186	0.310	0.400	0.592
.15	0.174	0.212	0.370	0.560
.20	0.142	0.182	0.380	0.578
.25	0.108	0.174	0.400	0.574
.30	0.136	0.206	0.402	0.574

A   Proofs of the Results

A.1   Theorem 1

The regression quantile process corresponding to ours is determined by the following optimization problem

$$\min_{\alpha,\beta} \sum_{t=1}^{n} \rho_\tau(y_t - \alpha - \beta D_t)$$

Let $\theta = (\alpha, \beta)'$, and $z_t = (1, D_t)'$, we can rewrite the optimization problem as

$$\widehat{\theta}(\tau) = \arg\min_{\theta} \sum_{i=1}^{n} \rho_\tau(y_t - z_t'\theta)$$

Notice that

$$y_t = \alpha + \beta D_t + w_t = \theta' z_t + w_t$$

$$Q_{y_t}(\tau|D_t) = \alpha(\tau) + \beta(\tau)D_t$$

and

$$\alpha(\tau) = \alpha + Q_v(\tau), \ \beta(\tau) = \beta + Q_u(\tau) - Q_v(\tau)\,.$$

24

Under regularity conditions and denoting $\psi_\tau(u) = \tau - I(u < 0)$, and $w_{t\tau} = y_t - z_t'\theta(\tau)$, we have

$$
\begin{aligned}
w_{t\tau} &= y_t - z_t'\theta(\tau) \\
&= y_t - \alpha(\tau) - \beta(\tau)D_t \\
&= \alpha + \beta D_t + v_t + (u_t - v_t)D_t - \alpha - Q_v(\tau) - [\beta + Q_u(\tau) - Q_v(\tau)] D_t \\
&= (v_t - Q_v(\tau))(1 - D_t) + (u_t - Q_u(\tau))D_t,
\end{aligned}
$$

and

$$
Q_{w_{t\tau}}(\tau|z_i) = 0
$$

where $Q_{w_{t\tau}}(\tau|z_i)$ is the τ-th conditional quantile of $w_{t\tau}$, and

$$
E\left[\psi_\tau(w_{t\tau})|x_t\right] = 0.
$$

Under our assumptions, the following Bahadur linear representation of $\widehat{\theta}(\tau)$ can be obtained:

$$
\sqrt{n}\left(\widehat{\theta}(\tau) - \theta(\tau)\right) = \left[\frac{1}{n}\sum_{t=1}^n f_t z_t z_t'\right]^{-1} \left(n^{-1/2}\sum_{t=1}^n z_t \psi_\tau(w_{t\tau})\right) + R_n
$$

where the reminder term R_n is $o_p(1)$ uniformly over τ. For any give τ,

$$
\sqrt{n}\left(\widehat{\theta}(\tau) - \theta(\tau)\right) \Rightarrow N\left(0, \tau(1-\tau)\Omega^{-1}H\Omega^{-1}\right)
$$

where

$$
f_t = f_t(F_{y_t}^{-1}(\tau|z_t)), \ \Omega = \lim \frac{1}{n}\sum_{t=1}^n f_t z_t z_t', \ H = \lim \frac{1}{n}\sum_{t=1}^n z_t z_t'.
$$

Under the null, $f_t(F_{y_t}^{-1}(\tau|z_t)) = f_u(F_u^{-1}(\tau)) = f_v(F_v^{-1}(\tau)) = f(F^{-1}(\tau))$, $\Omega = f(F^{-1}(\tau))\Omega_0$, and $\Omega_0 = H$. The limiting distribution of the regression quantile process is then given as follows

$$\sqrt{n}\left(\widehat{\theta}(\tau) - \theta(\tau)\right) = \frac{1}{f(F^{-1}(\tau))}\left[\frac{1}{n}\sum_{t=1}^{n}z_t z_t'\right]^{-1}\left(n^{-1/2}\sum_{t=1}^{n}z_t\psi_\tau(u_{t\tau})\right) + R_n$$

and

$$\sqrt{n}f(F^{-1}(\tau))\Omega_0^{1/2}\left(\widehat{\theta}(\tau) - \theta(\tau)\right) \Rightarrow \overline{W}_2(\tau), \text{ for } \tau \in \Pi$$

where $\overline{W}_2(\tau)$ is a two-dimensional Brownian bridge. In particular, at each τ,

$$\sqrt{n}\left(\widehat{\theta}(\tau) - \theta(\tau)\right) \Rightarrow N\left(0, \frac{\tau(1-\tau)}{f(F^{-1}(\tau))^2}\Omega_0^{-1}\right)$$

Thus, let $R = [0, 1]$, then

$$\sqrt{n}\left(\widehat{\beta}(\tau) - \beta(\tau)\right) \Rightarrow N\left(0, \frac{\tau(1-\tau)}{f(F^{-1}(\tau))^2}R\Omega_0^{-1}R'\right),$$

and the test statistic can be constructed based on

$$\widehat{V}_1(\tau) = \sqrt{n}\widehat{f}(F^{-1}(\tau))\left[R\Omega_0^{-1}R'\right]^{-1/2}\widehat{\beta}(\tau) \Rightarrow \overline{W}_1(\tau) \text{ for } \tau \in \Pi$$

where $\overline{W}_1(\tau)$ is a one-dimensional Brownian bridge.

References

Anderson, G., 1996, Nonparametric tests of stochastic dominance in income distributions, Econometrica, 64, 1183–1193.

Andrews, Donald W. K., 1993, Tests for Parameter Instability and Structural Change With Unknown Change Point, Econometrica, 61(4), 821-856.

Bai, Zhidong, Huixia Liu and Wing-Keung Wong, Enhancement of the Applicability of Markowitz's Portfolio Optimization by Utilizing Random Matrix Theory, Mathematical Finance, 19(4), 639-667.

Dardanoni, V., A. Forcina, 1998, A unified approach to likelihood inference on stochastic orderings in a nonparametric context, J. Am. Stat. Assoc., 93, 1112–1123.

Dardanoni, V., A. Forcina, 1999, Inference for Lorenz curve orderings, Econ. J. , 2, 49–75.

Davidson, R., and Jean-Yves Duclos, 2000, Statistical Inference for Stochastic Dominance and for the Measurement of Poverty and Inequality, Econometrica, 68(6), 1435-1464.

Egozcue, M., Wong, W.K., 2010, Gains from diversification: A majorization and stochastic dominance approach. European Journal of Operational Research 200, 893-900.

Jensen, M.C., 1969, Risk, the pricing of capital assets and the evaluation of investment portfolios, Journal of Business, 42, 167-247.

Kaur, A., B.L.S.P. Rao, H. Singh, 1994, Testing for second order stochastic dominance of two distributions, Econ. Theory 10, 849–866.

Klecan, L., R. McFadden, D. McFadden, 1991, A robust test for stochastic dominance, Working Paper, MIT & Cornerstone Research.

McFadden, D., 1989, Testing for stochastic dominance, in: T.B. Fomby, T.K. Seo (Eds.), Studies in the Economics of Uncertainty, Springer-Verlag, New York, pp. 113–132.

Markowitz, H.M., 1952, Portfolio selection, Journal of Finance, 7, 77-91.

Sharpe, W.F., 1964, Capital asset prices: Theory of market equilibrium under conditions of risk, Journal of Finance, 19, 452-442.

Tobin, J., 1958, Liquidity preference and behavior towards risk, Review of Economic Studies, 25, 65-86.

Treynor, J.L., 1965, How to rate management of investment funds, Harvard Business Review, 43, 63-75.

Tse, Y.K., X. Zhang, 2004, A monte carlo investigation of some tests for stochastic dominance, Journal of Statistical Computation & Simulation, 74, 361-378.