ESAIM: PS 21 (2017) 536–561
DOI: 10.1051/ps/2017015 **ESAIM: Probability and Statistics**
www.esaim-ps.org

VALIDATION OF POSITIVE EXPECTATION DEPENDENCE

Bogdan Ćmiel[1] and Teresa Ledwina[2]

Abstract. In this paper, we develop tests for positive expectation dependence. The proposed tests are based on weighted Kolmogorov−Smirnov type statistics. These originate from the function valued monotonic dependence function, describing local changes of the strength of the dependence. The resulting procedure is supported by a simple and insightful graphical device. This paper presents asymptotic and simulation results for such tests. We show that an inference relying on p-values and wild bootstrap allows to overcome inherent difficulties of this testing problem. Our simulations show that the new tests perform well in finite samples. A Danish fire insurance data set is examined to demonstrate the practical application of the proposed inference methods.

Mathematics Subject Classification. 60F05, 62G09, 62G10, 62G20.

Received January 9, 2017. Revised July 12, 2017. Accepted August 3, 2017.

1. Introduction

There is quickly growing evidence that the dependence structure of random vectors cannot be neglected in reliable data analysis. In particular, this problem is crucial in insurance and finance. For example, Dhaene and Goovaerts [16] presented several sources of dependencies between various risks and searched for the type of dependency between individuals that gives rise to the largest stop-loss premiums. They also summarized the work of Kaas [25], who demonstrated that such dependencies may have disastrous effects on stop-loss premiums. Albers [1] showed that substantial deviations in the fair price of stop-loss premiums may occur even when there are small departures from independence. Denuit and Scaillet [15] emphasized the following: "In the management of large portfolios, the main risk is the joint occurrence of a number of default events or the simultaneous downside evolution of prices. A better knowledge of the dependence between financial assets or claims is crucial to assess the risk of loss clustering." Extensive work by practitioners has also revealed that some well known global measures of the strength of dependence, for example the correlation coefficient, are not capable of explaining the complex character of many relations, while some other existing or new notions may be much more useful in current practice and specific applications. For some illustration, discussion and new ideas, see Embrechts *et al.* [18], Denuit and Scaillet [15], Baur [4], Berentsen and Tjøstheim [6], Li *et al.* [35], Ledwina [33] and the references therein.

Keywords and phrases. Hypothesis testing, expectation dependence, Lorenz curve, monotonic dependence function, multiplier central limit theorem, wild bootstrap, Zenga curve.

[1] AGH University of Science and Technology, Faculty of Applied Mathematics, Al. Mickiewicza 30, 30-059 Cracow, Poland. cmielbog@gmail.com

[2] Polish Academy of Sciences, Institute of Mathematics, ul. Kopernika 18, 51-617 Wrocław, Poland. ledwina@impan.pl

In recent years, the role of the notion of positive expectation dependence has become increasingly important in various research areas. This notion was introduced formally by Yanagimoto [46] and investigated thoroughly by Kowalczyk and Pleszczyńska [29], who studied a function valued measure of the strength of dependence. Later, this approach was rediscovered in econometrics by Wright [45]. Interest in this concept has grown markedly in the few last years, due to its fruitful applications in finance. Below, we provide the definition of this concept, recall some alternative names for it which have appeared in the literature, and briefly mention a few applications.

Let (X,Y) be a bivariate random variable with finite expectations and continuous marginal distribution functions. Positive expectation dependence of Y on X occurs when

$$E(Y|X \le x) \le EY \quad \text{for all} \quad x \in \mathbb{R}, \tag{1.1}$$

or equivalently $E(Y|X > x) \ge EY$ for all $x \in \mathbb{R}$. The interpretation of this notion is clear. The information that X is greater than some value x increases the expected value of Y. This notion is weaker than the well known concept of positive quadrant dependence (PQD in short). Positive expectation dependence and positive quadrant dependence are two formalizations of so-called monotonic dependence. Roughly speaking, Y is monotone dependent on X if large values of Y tend to be associated with large values of X (positive dependence) or with small values of X (negative dependence).

Some earlier notation and alternative names for (1.1) are: $\mathcal{P}(1,E)$ (Yanagimoto [46]; see also Yanagimoto and Shibuya [47] for the corresponding definition of $\mathcal{R}_E$), EQD^+ (Kowalczyk and Pleszczyńska [29]), positive expectation dependence (Wright [45]), $PQDE$ and positive quadrant dependence in expectation (Balakrishnan and Lai [2]), as well as first-degree positive expectation dependence (Hong *et al.* [24]). We shall follow the terminology introduced by Wright and the notation EQD^+ proposed by Kowalczyk and Pleszczyńska.

Ćwik *et al.* [10] applied this notion to the chronological ordering of Poisson streams. Wright [45] showed its usefulness in solving a classical problem of portfolio theory. Hong *et al.* [24] used the concept of expectation dependence to provide necessary and sufficient conditions for Mossin's Theorem to hold when initial wealth is random. Egozuce *et al.* [17] demonstrated that, under EQD^+, Grüss-type inequalities for covariances can be considerably sharpened and concluded that the notion of expectation dependence is ideally suited for analyzing the covariance between Y and certain non-decreasing functions of X. They also mentioned that, in a number of situations, analyzing the allocation of capital involves such quantities. Guo *et al.* [22] extended the concept of expectation dependence and applied it to their generalization of the Rothschild−Stiglitz type of increasing risk in the framework of background risk. Denuit and Mesfioui [14] provided some further insight into this problem. Moreover, recent work by Denuit *et al.* [13] and Guo and Li [21] gave an excellent discussion of a wide range of further applications of (1.1). Guo and Li [21] concluded that expectation dependence is a key concept in many economic and financial studies.

The above discussion shows that there is strong motivation to provide tools to test for the existence of positive expectation dependence. The first test for the existence of the relation (1.1) was recently proposed by Zhu *et al.* [49]. Their approach is to estimate both sides of (1.1) and to reject the hypothesis of positive expectation dependence when the supremum (over the values of x) of the estimates of the differences of the form $E(Y|X \le x) - EY$ is too large. This approach was recently advanced by Guo and Li [21], who provided confidence bounds for the curve $ED(x) = EY - E(Y|X \le x)$.

We propose an alternative test of (1.1). Our starting point is the simple observation that one can re-parametrize (1.1) by introducing $x_p = F^{-1}(p)$, where $p \in (0,1)$, F denotes the marginal distribution function of X and F^{-1} denotes its inverse. Using this notation, (1.1) is equivalent to

$$E(Y|X \le x_p) \le EY \quad \text{for all} \quad p \in (0,1). \tag{1.2}$$

In other words, in contrast to using a fixed threshold based on (1.1), we shall consider a threshold based on quantile hits. Quantile hits have a well established position in the statistical literature, especially in economics and in social sciences, and sometimes are simply indispensable to making meaningful inference. For some discussion of this issue, see Handcock and Morris [23], as well as Linton and Whang [36]. Moreover, this re-parametrization allows us to make additional interpretations regarding the form of expectation dependence. To be specific, observe that, after elementary integration in (1.2), we can present an equivalent formulation of (1.1) as follows:

$$GLC(p) \le p \quad \text{for all} \quad p \in (0,1), \quad \text{where} \quad GLC(p) = \frac{1}{EY}\int_0^p E(Y|X = x_s)\mathrm{d}s,$$

provided that $EY \neq 0$. The function $GLC(p)$ is called the generalized Lorenz curve for the regression function $E(Y|X)$. Typically, the definition of $GLC(p)$ is restricted to non-negative random variables X and Y. For $Y = X$, $GLC(p)$ is the standard Lorenz curve, denoted here by $LC(p)$, while in the general case it can be interpreted (at least in some important cases) as the Lorenz curve in the presence of a covariate. For more details on generalized Lorenz curves, see Muliere and Petrone [37]. Zenga [48] introduced a new curve describing economic inequality, denoted here by $ZC(p)$, given by $ZC(p) = 1 - E(Y|Y \le y_p)/E(Y|Y > y_p),\ p \in (0,1)$, for a non-negative Y with continuous distribution function and finite expectation. For more details, see Nair and Sreelakshmi [38]. Analogously to $GLC(p)$, we introduce $GZC(p) = 1 - E(Y|X \le x_p)/E(Y|X > x_p)$ and note that, for non-negative Y, (1.1) holds if and only if $GZC(p) \ge 0$ for all $p \in (0,1)$. Note also that in Rao and Zhao [39] the function $m(s) = E(Y|X = x_s)$ is called the quantile regression function of Y on X, while the integral $f(p) = \int_0^p m(s)\mathrm{d}s$ is called the cumulative quantile regression function. In contrast, the integral $f(p)$ is called the correlation function of Y on X in Kowalczyk and Szczesny [30], where further discussion and references can also be found. The function $f(p)$ was also studied, interpreted in terms of the absolute concentration curve and its role in economics extensively discussed in Schechtman *et al.* [42].

It is intuitively clear that, when making inference regarding (1.1) or (1.2), it is useful not only to look at the absolute differences between the two sides of these inequalities, but also to take into account their relative magnitudes. For this purpose, we consider a standardized difference based on (1.2). This standardization is closely related to the monotonic dependence function introduced and investigated in Kowalczyk and Pleszczyńska [29] and later papers. This dependence function possesses several appealing properties and allows us, in particular, to measure and visualize the strength of the relationship (1.2). The form of this standardization depends on whether or not (1.2) holds. If (1.2) is satisfied, then the monotonic dependence function $\mu_{Y,X}(p),\ p \in (0,1)$, takes the form

$$\mu_{Y,X}(p) = \mu^+_{Y,X}(p) = \frac{E(Y|X > x_p) - EY}{E(Y|Y > y_p) - EY} = \frac{EY - E(Y|X \le x_p)}{EY - E(Y|Y \le y_p)} = \frac{p - GLC(p)}{p - LC(p)}, \tag{1.3}$$

with $y_p = G^{-1}(p)$, where G^{-1} is the inverse of G, the distribution function of Y. In contrast, when (1.2) is violated for some $p \in (0,1)$, then $GLC(p) > p$ and the function $\mu_{Y,X}$ takes the form

$$\mu_{Y,X}(p) = \mu^-_{Y,X}(p) = -\mu^+_{-Y,X}(p) = \frac{E(Y|X > x_p) - EY}{EY - E(Y|Y \le y_{1-p})} = \frac{p - GLC(p)}{(1-p) - LC(1-p)}\cdot \tag{1.4}$$

Analogous expressions involving $ZC(p)$ and $GZC(p)$ can be obtained in a similar manner. A concise formula for $\mu_{Y,X}(p)$ is given in Section 2.1.

Note that an equivalent form of (1.2) is given by $E(Y|X > x_p) \ge EY$ for all $p \in (0,1)$, since the denominator in (1.4) is positive for all $p \in (0,1)$. Therefore, it is natural to reject positive expectation dependence when the infimum (over the values of p) of the estimates of $\mu^-_{Y,X}(p)$ is sufficiently small. Observe also that the functions in both the numerator and the denominator of (1.3) and (1.4) tend to 0 as p tends to either 0 or 1. Therefore, the weight used in (1.4) blows up the value of the numerator of (1.4) when p is close to 0 or 1. In our simulation study, we also consider lighter weight. For the details, see Section 4. Note that if one would like to verify the negative expectation dependence of Y on X, then the supremum (over the values of p) of the estimate of $\mu^+_{Y,X}(p)$ would be a good candidate for the appropriate test statistic.

From the form of (1.4), it is clear that when estimating $\mu^-_{Y,X}(p)$ and the dependence function, in general, we are faced by the analysis of processes that naturally appear in studies of generalized and classical Lorenz curves. Hence, we shall rely on the approach and asymptotic results of Davydov and Egorov [11]. For alternative approaches, see Goldi [19], Rao and Zhao ([39, 40]), Tse [43] and the references therein. The limiting processes of the natural estimators of the numerator and the denominator in (1.4) are Gaussian and their parameters depend on the underlying distribution of (X, Y). Hence, when constructing the appropriate rejection region, we have to overcome two basic difficulties: the fact that the parameters of the limiting distributions of the underlying processes are unknown and the need to control the inequality constraints under the null hypothesis. To solve both problems, we shall follow the ideas developed by Barrett and Donald [3] in the context of testing for stochastic dominance. In particular, we shall apply a rejection rule relying on p-values and combine it with a resampling procedure based on multiplier central limit theory in $C[0, 1]$ space. The resampling plan belongs to a class of wild bootstrap procedures. The details are given in Section 3 and Appendix A.

The remainder of this paper is structured as follows. In Section 2.1 we recall the definition of the monotonic dependence function and briefly summarize its properties. When one is investigating the relation (1.2) for some given model, the monotonic dependence function enables seeing whether (1.2) holds, for which values of p (1.2) is possibly violated, and what is the local (at a given value of p) strength of the monotonic dependence. In Section 2.2 we present a natural estimator of this function, introduced and studied in Kowalczyk [26]. This estimator inherits all the basic properties of the monotonic dependence function. It reflects the magnitude of and local changes in the strength of the expectation dependence. Therefore, this estimator can be used as a diagnostic tool to observe, understand, and infer possible reasons for any departure from the null hypothesis. Its shape can also be useful when selecting an appropriate model for the data at hand, or at least to invalidate several inappropriate candidates for models. In short, the estimated monotonic dependence function plays a similar role in verifying EQD^+ to the one played by a $Q - Q$ plot in classical goodness-of-fit tests. Obviously, it would also be desirable if such analysis could be accompanied by the construction of confidence bounds for the unknown monotonic dependence function. This, however, requires further work on the corresponding asymptotic theory and is beyond the scope of the present study. The paper by Csörgő *et al.* [9], concerning confidence intervals for the classical Lorenz curve, illustrates the inherent complexity of the underlying problem well. Section 3 presents a description of our basic test procedure and collates results on its asymptotic behavior. The corresponding proofs are given in Appendix A. In Section 4 we present the empirical behavior of the test based on our basic construction, defined in Section 3, along with analogous results for a variant of this test. Moreover, these two solutions are compared with the test introduced in Zhu *et al.* [49]. In this Section we display the results mainly graphically, while the full set of numerical results is presented in Appendix C. Our simulation study was extensive and concerned twelve different models of various monotonic dependence functions. In approximately half of the cases considered, the solution of Zhu *et al.* [49] is slightly better than ours. However, the remaining cases clearly illustrate the positive effect of the weighting applied in our construction. The average powers of our tests are several percent greater than that of the test presented by Zhu *et al.* [49]. On the basis of our simulation experiments, we have also proposed a third test, whose average power dominates the power of the tests considered so far. Analytical formulas for the monotonic dependence function for the models considered in Section 4 are presented in Appendix B. In Section 5 we illustrate our approach by applying it to a well known data set. Section 6 provides some concluding remarks.

2. Monotonic dependence function

2.1. Definition and properties

Throughout this paper we assume that (X, Y) belongs to the set $\mathcal{B}$ of random vectors which have joint distribution function $H(x, y)$, finite expectations and continuous marginal distribution functions $F(x)$ and $G(y)$, respectively.

Set

$$l_{Y,X}(p) = E\big\{Y\big[p - I(X \leq x_p)\big]\big\}$$

and observe that (1.2) holds if and only if $l_{Y,X}(p) \geq 0$ for all $p \in (0, 1)$.

For any $(X,Y) \in \mathcal{B}$, define the function $\mu_{Y,X} : (0,1) \to [-1,1]$, called the monotonic dependence function of Y on X, as follows

$$\mu_{Y,X}(p) = \mu^+_{Y,X}(p) \quad \text{if} \quad l_{Y,X}(p) \geq 0$$

and

$$\mu_{Y,X}(p) = \mu^-_{Y,X}(p) \quad \text{if} \quad l_{Y,X}(p) < 0,$$

where $\mu^+_{Y,X}(p)$ and $\mu^-_{Y,X}(p)$ are given by (1.3) and (1.4), respectively.

Before we list some of the properties of $\mu_{Y,X}(p)$, let us introduce some conventions. Namely, we shall say that the function $h : \mathbb{R} \to \mathbb{R}$ is F-increasing (decreasing), if for all $s,t \in \mathbb{R}$ the relation $F(s) < F(t)$ implies that $h(s) < h(t)$ $(h(s) > h(t))$. Moreover, for any two random variables U and V, we write $U \doteq V$, if $\Pr(U = V) = 1$. Finally, by $U =^d V$ we mean that U and V have the same distributions.

Lemma 2.1. *If $(X,Y) \in \mathcal{B}$ and $p \in (0,1)$, then*

1. $\mu_{Y,X}(p)$ *is continuous.*
2. *For any real a and b, $a \neq 0$, it holds that $\mu_{aY+b,f(X)}(p) = (\operatorname{sgn} a)\mu_{Y,X}(p)$ if f is F-increasing and $\mu_{aY+b,f(X)}(p) = (-\operatorname{sgn} a)\mu_{Y,X}(1-p)$ if f is F-decreasing.*
3. $-1 \leq \mu_{Y,X}(p) \leq 1$.
4. $\mu_{Y,X}(p) = 1 \Leftrightarrow \Pr(X < x_p, Y > y_p) = \Pr(X > x_p, Y < y_p) = 0$. *Moreover, $\mu_{Y,X}(p) \equiv 1 \Leftrightarrow$ there exists an F-increasing f such that $Y \doteq f(X) \Leftrightarrow H(x,y) = \min\{F(x), G(y)\}$.*
5. $\mu_{Y,X}(p) = -1 \Leftrightarrow \Pr(X < x_{1-p}, Y < y_p) = \Pr(X > x_{1-p}, Y > y_p) = 0$. *Moreover, $\mu_{Y,X}(p) \equiv -1 \Leftrightarrow$ there exists an F-decreasing f such that $Y \doteq f(X) \Leftrightarrow H(x,y) = \max\{0, F(x) + G(y) - 1\}$.*
6. $\mu_{Y,X}(p) \equiv 0 \Leftrightarrow E(Y|X) \doteq EY$.
7. *Suppose that there exists an increasing function $h : \mathbb{R} \to \mathbb{R}$ and ρ, $0 \neq \rho \in [-1,1]$, for which $h(X)$ and $(sgn\ \rho)Y$ have the same distribution. Then $\mu_{Y,X}(p) \equiv \rho$ if and only if $E(Y|X) \doteq \rho h(X) + (1-|\rho|)EY$. Moreover, if the correlation coefficient of X and Y exists and is equal to $\rho \neq 0$, then $\mu_{Y,X}(p) \equiv \rho$ if and only if $E(Y|X) \doteq \rho h(X) + (1-|\rho|)EY$ and h is linear.*
8. *Suppose that (X,Y) and (X',Y') have identical pairs of marginal distributions F and G. The equality $\mu_{Y,X} \equiv \mu_{Y',X'}$ holds if and only if $E(Y|X)$ and $E(Y'|X')$ have the same distribution. This shows that the form of $\mu_{Y,X}$ depends only on the distribution of $E(Y|X)$.*
9. *Consider $(X,Y) \in \mathcal{B}$ such that $F(x) = G(x)$ for all x and F possesses a strictly positive density. Then*
   - *if $r(x) = E(Y|X=x)$ is continuous, nonlinear, non-decreasing and convex (concave), then $\mu_{X,Y}(p)$ is positive and increasing (decreasing). If $r(x) = E(Y|X=x)$ is continuous, nonlinear, non-increasing and convex (concave), then $\mu_{X,Y}(p)$ is negative and increasing (decreasing).*
   - *$r(x)$ is linearly increasing $\Leftrightarrow$ $\mu_{Y,X}(p)$ is constant and positive. Moreover, $r(x)$ is linearly decreasing $\Leftrightarrow$ $\mu_{Y,X}(p)$ is constant and negative.*
   - *$r(x)$ is constant $\Leftrightarrow$ $\mu_{Y,X}(p) \equiv 0$.*

Properties 1−6 and 8 were stated and proved in Kowalczyk and Pleszczyńska [29], while property 7 comes from Kowalczyk [26]. These properties are also valid without the assumption on the continuity of the marginal distributions F and G. This was shown in Kowalczyk [26]. Property 9 is a special case of Theorem 1 and Corollary 1 in Kowalczyk [27], which was designed to studying a non-trivial interpretation of the information on the joint distribution H contained in the shape of the monotonic dependence function $\mu_{Y,X}$. Some preliminary results on the shape of the monotonic dependence function can be found in Ćwik *et al.* [10].

Remark 2.2. Suppose that (X,Y) and (X',Y') are from $\mathcal{B}$ and have identical pairs of marginal distributions. Then $\mu_{Y,X}(p) \geq \mu_{Y',X'}(p)$ for all $p \in (0,1) \Leftrightarrow E(Y|X \leq x_p) \leq E(Y'|X' \leq x_p)$ for all $p \in (0,1)$. In words, if small values of Y are more strongly associated with small values of X than small values of Y' are associated with small values of X', then the dependence function of (X,Y) takes a larger value than that of (X',Y').

From property 3, this tendency is strongest when $\mu_{Y,X}(p) \equiv 1$ and weakest when $\mu_{Y,X}(p) \equiv -1$, *i.e.* when there exist monotonic functions that relate Y to X. We can also say that the tendency is undetermined if $E(Y|X \leq x_p) = EY$ for each p or, equivalently, if $\mu_{Y,X}(p) \equiv 0$. In the case of EQD^+ distributions, the weakest form of dependence corresponds to $\mu_{Y,X}(p) \equiv 0$, *i.e.* the function describing the regression of Y on X is constant; *cf.* point 6 of Lemma 2.1. From (1.3), we can rephrase this as follows: larger values of $\mu_{Y,X}$ mean a greater similarity between the generalized Lorenz curve for $r(X) = E(Y|X)$ and the Lorenz curve for Y. Similar conclusions hold in the case of negative expectation dependence.

From the above, $\mu_{Y,X}(p) \equiv \mu_{Y',X'}(p) \Leftrightarrow E(Y|X) =^d E(Y'|X') \Leftrightarrow$ the GLCs of (X,Y) and (X',Y') coincide.

2.2. Empirical monotonic dependence function

As mentioned earlier, our test statistic is based on an estimator of $\mu_{Y,X}$. Therefore, we start with a natural estimator $\hat{\mu}_{Y,X}$ of the function $\mu_{Y,X}$, which was proposed and investigated in Kowalczyk [26]. That paper extended the definition of $\mu_{Y,X}$ to all random vectors (X,Y) which have finite expectations and marginal distributions that are not concentrated on one point. This extension preserves the properties 1–8 of $\mu_{Y,X}$ that were listed in Section 2.1. The estimator $\hat{\mu}_{Y,X}$ is simply the monotonic dependence function calculated for the empirical distribution of the sample.

To introduce this estimator, we need some notation. Consider $(X,Y) \in \mathcal{B}$ and independent, identically distributed random vectors $(X_1,Y_1),\ldots,(X_n,Y_n)$ with the same distribution as (X,Y). Let $X_{(i)}$ and $Y_{(i)}, i = 1,\ldots,n$, be the i-th order statistics in the sequences $X_1,\ldots,X_n$ and $Y_1,\ldots,Y_n$, respectively. The p-th sample quantiles $\hat{x}_p$ and $\hat{y}_p$ are defined to be $\hat{x}_p = X_{(\lfloor np \rfloor+1)}$ and $\hat{y}_p = Y_{(\lfloor np \rfloor+1)}$, accordingly, where $\lfloor \bullet \rfloor$ denotes the integer part of the real number $\bullet$. From the above, the estimator $\hat{\mu}_{Y,X}$ of $\mu_{Y,X}$ is of the form

$$\hat{\mu}_{Y,X}(p) = \hat{\mu}^+_{Y,X}(p) \quad \text{if} \quad L_{Y,X}(p) \geq 0$$

and

$$\hat{\mu}_{Y,X}(p) = \hat{\mu}^-_{Y,X}(p) \quad \text{if} \quad L_{Y,X}(p) < 0,$$

while

$$\hat{\mu}^+_{Y,X}(p) = \frac{L_{Y,X}(p)}{M^+_Y(p)}, \quad \hat{\mu}^-_{Y,X}(p) = \frac{L_{Y,X}(p)}{M^-_Y(p)},$$

$$L_{Y,X}(p) = \frac{1}{n}\sum_{i=1}^{n} Y_i\big[p - I(X_i < \hat{x}_p) - (np - \lfloor np \rfloor)I(X_i = \hat{x}_p)\big],$$

$$M^+_Y(p) = \frac{1}{n}\sum_{i=1}^{n} Y_i\big[p - I(Y_i < \hat{y}_p)\big] - \big(p - \frac{1}{n}\lfloor np \rfloor\big)\hat{y}_p,$$

$$M^-_Y(p) = \frac{1}{n}\sum_{i=1}^{n} Y_i\big[I(Y_i > \hat{y}_{1-p}) - p\big] + \big(p - 1 + \frac{1}{n}\lfloor n(1-p)+1 \rfloor\big)\hat{y}_{1-p}$$

where $I(A)$ denotes the indicator function of the set A. Observe that, given any sample, the functions $L_{Y,X}(p)$, $M^+_Y(p)$ and $M^-_Y(p)$ are continuous on (0,1). Moreover, $M^-_Y(p) = M^+_{-Y}(p)$. Also, $M^+_Y(p)$ and $M^-_Y(p)$ are positive on (0,1). For $(X,Y) \in \mathcal{B}$, the estimator $\hat{\mu}_{Y,X}$ is consistent; *cf.* Kowalczyk [26].

3. Construction of the test and its properties

Since our testing problem can also be rephrased as

$$\mathcal{H}_0 : \mu_{Y,X}(p) \geq 0 \quad \text{for all} \quad p \in (0,1),$$

against

$$\mathcal{H}_1 : \mu_{Y,X}(p) < 0 \quad \text{for some} \quad p \in (0,1),$$

the following is a natural decision rule: reject $\mathcal{H}_0$ if $\inf_{0<p<1} \hat{\mu}_{Y,X}(p)$ is sufficiently small or, equivalently, when $\sup_{0<p<1}\{-\hat{\mu}_{Y,X}(p)\}$ is sufficiently large. However, from Proposition 3.2, the standardized estimators $\hat{\mu}^+_{Y,X}(p)$ and $\hat{\mu}^-_{Y,X}(p)$ are asymptotically centered Gaussian processes. Therefore, for large samples, when testing at a significance level from (0,1/2), the part $\hat{\mu}^+_{Y,X}(p)$ of $\hat{\mu}_{Y,X}(p)$ leads to rejection of the null hypothesis negligibly often. Hence, a simpler procedure, which rejects $\mathcal{H}_0$ for sufficiently large values of $\sup_{0<p<1}\{-\hat{\mu}^-_{Y,X}(p)\}$, is equally good when standard significance levels are used.

In what follows, we restrict attention to values of p in $[\epsilon, 1-\epsilon]$, where $\epsilon \in (0, 1/2)$. Remembering that the denominator of $\mu^-_{Y,X}(p)$ and, in consequence its estimator as well, tend to 0 for p close to 0 or 1, introducing such a sub-interval results in our solution being more stable for finite samples. We discuss this question further in Section 4. In consequence, given $\epsilon \in (0, 1/2)$, we set

$$S^-(\epsilon) = \sup_{\epsilon \le p \le 1-\epsilon} \sqrt{n}\{-\hat{\mu}^-_{Y,X}(p)\} = \sup_{\epsilon \le p \le 1-\epsilon} \left\{ \frac{-\sqrt{n}L_{Y,X}(p)}{M^-_Y(p)} \right\}$$

and reject $\mathcal{H}_0$ for large positive values of $S^-(\epsilon)$. To implement this rule in practice, we shall rely on some asymptotic results and Monte Carlo methods, which mimic the asymptotic results for finite samples. We present these issues in the two next subsections. Here, we conclude with a useful observation on an alternative form of $S^-(\epsilon)$.

Remark 3.1. Note that, given realizations of $(X_1, X_2), \ldots, (X_n, Y_n)$, both the numerator and the denominator of $\hat{\mu}_{Y,X}(p)$ are continuous functions of p, $p \in (0, 1)$, that are linear on the intervals $([k-1]/n, k/n),\ k = 1, \ldots, n$. Therefore, their ratio is piecewise monotonic on these intervals. In particular, this implies that

$$\sup_{\epsilon \le p \le 1-\epsilon} \{-\hat{\mu}^-_{Y,X}(p)\} = \max_{k(\epsilon) \le j \le n-k(\epsilon)} \left\{ \frac{-L_{Y,X}(j/n)}{M^-_Y(j/n)} \right\},$$

where $k(\epsilon) = \lfloor n\epsilon \rfloor$. A similar expression is valid for $\sup_{\epsilon \le p \le 1-\epsilon} \hat{\mu}^+_{Y,X}(p)$.

3.1. Auxiliary asymptotic results

We introduce the two continuous functions on $(0, 1)$

$$m^+_Y(p) = E\{Y[p - I(Y \le y_p)]\}, \quad m^-_Y(p) = E\{Y[(1 - p - I(Y \le y_{1-p})]\},$$

and the processes

$$Z_n(p) = \sqrt{n}\Big\{L_{Y,X}(p) - l_{Y,X}(p)\Big\}, \qquad V^+_n(p) = \frac{Z_n(p)}{M^+_Y(p)}, \quad V^-_n(p) = \frac{Z_n(p)}{M^-_Y(p)}, \quad p \in (0,1),$$

which take values in the space $C[0, 1]$ of continuous functions on $[0, 1]$. By $\Rightarrow$ we denote weak convergence in the space of functions on $[0, 1]$ under consideration. The uniform distance is applied in all function spaces under consideration and $\xrightarrow{\Pr}$ stands for convergence in probability. Although some of the results presented below could be stated for $C[0, 1]$, in view of the form of our test statistic, we restrict our attention throughout to $[\epsilon, 1-\epsilon]$. This means that fewer assumptions are required and a more concise presentation is possible.

Proposition 3.2. *Let ϵ be a constant in $(0, 1/2)$. Suppose $EY^2 < \infty$ and the quantiles of Y are uniquely determined. Then*

$$Z_n \Rightarrow Z, \;\; M^+_Y \xrightarrow{\Pr} m^+_Y, \;\; M^-_Y \xrightarrow{\Pr} m^-_Y, \;\; V^+_n \Rightarrow V^+, \;\; \textit{and} \;\; V^-_n \Rightarrow V^-, \;\; \textit{in} \;\; C[\epsilon, 1-\epsilon], \tag{3.1}$$

where Z, V^+ and V^- are centered Gaussian processes concentrated on $C[\epsilon, 1-\epsilon]$.

A justification of (3.1), the structure of the limiting processes and their parameters are briefly discussed in Appendix A. Basically, these results follow from Theorem 1 in Davydov and Egorov [11]. The covariance function of V^- is not simple and depends on the underlying distribution function H in a rather involved way. However, Proposition 3.2 allows us to formulate an asymptotic result for a prototype version of our final test. This is the subject of Proposition 3.3, below.

To formulate our first result on $S^-(\epsilon)$, define $\bar{V}^-(\epsilon)$ such that $\bar{V}^-(\epsilon) = \sup_{\epsilon \le p \le 1-\epsilon}\{-V^-(p)\}$.

Proposition 3.3. *Given $\epsilon \in (0, 1/2)$ and the assumptions of Proposition 3.2, for any $c \in (0, \infty)$ it holds that*

(i) *if $\mathcal{H}_0$ is true,*

$$\limsup_{n\to\infty} \Pr(S^-(\epsilon) > c) \le \Pr(\bar{V}^-(\epsilon) > c) \tag{3.2}$$

with equality when $l_{Y,X}(p) \equiv 0$ on $[\epsilon, 1-\epsilon]$;

(ii) *if $\mathcal{H}_0$ is false,*

$$\lim_{n\to\infty} \Pr(S^-(\epsilon) > c) = 1,$$

provided that there exists a $p_0 \in [\epsilon, 1-\epsilon]$ such that $l_{Y,X}(p_0) < 0$.

Note that the distribution of the random variable $\bar{V}^-(\epsilon)$ dominates the limiting law of $S^-(\epsilon)$ and the bound given by (3.2) is exact for those elements of $\mathcal{H}_0$ for which $\mu_{Y,X}(p) \equiv 0$ on $[\epsilon, 1-\epsilon]$. Set $\alpha_H(c) = \Pr(\bar{V}^-(\epsilon) > c)$. The above proposition implies that in the case when, given $\alpha \in (0, 1/2)$, one can find a c such that $\alpha_H(c) = \alpha$, the test rejecting $\mathcal{H}_0$ when $S(\epsilon) > c$ has the significance level α for the one-sided composite null hypothesis $\mathcal{H}_0$. However, the bounding random variable is not observable. In the next step, we describe a resampling method which allows us to approximate the distribution of $\bar{V}^-(\epsilon)$.

3.2. Multiplier method, working test procedure and main asymptotic result

Let $U = (U_1, \dots, U_n)$ denote a sample of independent $\mathcal{N}(0, 1)$ random variables that are independent of the sample $(X_1, Y_1), \dots, (X_n, Y_n)$. Define

$$L^U_{Y,X}(p) = \frac{1}{n}\sum_{i=1}^{n} U_i\Big\{(Y_i - \bar{Y})\big[p - I(X_i < \hat{x}_p) - (np - \lfloor np \rfloor) I(X_i = \hat{x}_p)\big] - L_{Y,X}(p)\Big\},$$

where $\bar{Y} = \sum_{i=1}^{n} Y_i/n$. Denote by $\hat{\mu}^{U-}_{Y,X}(p)$ the randomized estimator of $\mu^-_{Y,X}(p)$ defined as

$$\hat{\mu}^{U-}_{Y,X}(p) = \frac{L^U_{Y,X}(p)}{M^-_Y(p)}$$

and set

$$S^{U-}(\epsilon) = \sup_{\epsilon \le p \le 1-\epsilon} \Big\{ -\sqrt{n}\hat{\mu}^{U-}_{Y,X}(p)\Big\} \tag{3.3}$$

for the randomized version of $S^-(\epsilon)$.

Given the sample $(X_1, Y_1), \dots, (X_n, Y_n)$, consider the corresponding value of $S^-(\epsilon)$ and the probability

$$\hat{p}(n, \epsilon) = \Pr{}^U\left(\sup_{\epsilon \le p \le 1-\epsilon}\Big\{-\sqrt{n}\hat{\mu}^{U-}_{Y,X}(p)\Big\} > S^-(\epsilon)\right) = \Pr{}^U\Big(S^{U-}(\epsilon) > S^-(\epsilon)\Big). \tag{3.4}$$

The probability defined by (3.4) depends on the normal random variables $U_1, \dots, U_n$ and is conditional on the realizations of (X, Y)'s. To emphasize this, this probability is denoted by $\Pr^U$. The following is an operational variant of the test with the critical region $\{S^-(\epsilon) > c\}$:

$$\text{reject} \quad \mathcal{H}_0 \quad \text{if} \quad \hat{p}(n, \epsilon) < \alpha. \tag{3.5}$$

Theorem 3.4. *Suppose that the assumptions of Proposition 3.3 are satisfied and $\alpha \in (0, 1/2)$. Then*

(i) *if $\mathcal{H}_0$ is true,*

$$\limsup_{n\to\infty} \Pr\big(\hat{p}(n,\epsilon) < \alpha\big) \le \alpha \tag{3.6}$$

with equality when $l_{Y,X}(p) \equiv 0$ on $[\epsilon, 1-\epsilon]$;

(ii) *if $\mathcal{H}_0$ is false,*

$$\lim_{n\to\infty} \Pr\big(\hat{p}(n,\epsilon) < \alpha\big) = 1, \tag{3.7}$$

provided that there exists a $p_0 \in [\epsilon, 1-\epsilon]$ such that $l_{Y,X}(p_0) < 0$.

Remark 3.5. In order to implement (3.5), one has to calculate the supremum in (3.4) and to approximate the quantity $\hat{p}(n,\epsilon)$, which is not directly observable. The first question is resolved by a similar argument to the one made in Remark 3.1. This shows that the supremum can be calculated easily and explicitly. The second question is solved by standard simulation techniques. More specifically, let $U^{(1)}, \ldots, U^{(r)}$ denote independent copies of $U = (U_1, \ldots, U_n)$. Then the probability $\hat{p}(n,\epsilon)$ is approximated using

$$\hat{p}(n,\epsilon) \simeq \frac{1}{r}\sum_{i=1}^{r} I\Big(S^{U^{(i)}-} > S^-(\epsilon)\Big).$$

4. Simulation study

4.1. Description of the experiments, some motivations and discussion

In the Monte Carlo experiments, we investigated the finite sample properties of $S^-(\epsilon) = \sup_{\epsilon\le p\le 1-\epsilon}\{-\sqrt{n}L_{Y,X}(p)/M_Y^-(p)\}$ and its re-weighted variant given by

$$R^-(\epsilon) = \sup_{\epsilon\le p\le 1-\epsilon}\left\{\frac{-\sqrt{n}L_{Y,X}(p)}{\sqrt{M_Y^-(p)}}\right\}.$$

The statistic $R^-(\epsilon)$ puts a lighter weight on extreme values of Y than $S^-(\epsilon)$ does. This modification does not imply the need for major modifications in the analysis of the asymptotics. Therefore, we do not present obvious analogies of our basic asymptotic results concerning $S^-(\epsilon)$. Some more general weight functions could be considered as well. However, in this paper we only present an analysis of the natural weighting function M_Y^- and its simple variant.

We also compare $S^-(\epsilon)$ and $R^-(\epsilon)$ to the unweighted test statistic of Zhu *et al.* [49], who reject $\mathcal{H}_0$ for large values of

$$T = \sup_{x\in\mathbb{R}}\left\{-\frac{1}{\sqrt{n}}\sum_{i=1}^{n}(Y_i - \bar{Y})\left[I(X_i > x) - \frac{1}{n}\sum_{i=1}^{n} I(X_i > x)\right]\right\}.$$

Note that the process defined by T has jumps at successive order statistics $X_{(1)}, \ldots, X_{(n)}$. The process $L_{Y,X}(p)$, considered here, has continuous paths, but also attains local maxima at the order statistics. Moreover, $L_{Y,X}(p)$ is invariant to a shift of the Y_i's. Therefore, in principle, we can interpret our test statistics as weighted variants of T.

The procedure based on estimating p-values using the multiplier method was applied for each of these three statistics.

We designed the simulation scheme according to Zhu *et al.* [49]. Namely, $r = 2000$ Monte Carlo replications are used to approximate p-values, and each experiment is repeated 1000 times to compare the empirical significance level and power of these three tests at a nominal significance level of $\alpha = 0.05$.

In our study, we consider the two linear regression models introduced in Zhu *et al.* [49] and ten additional ones, which include quadratic regression, heavy tailed distributions, a bivariate exponential model, mixtures of bivariate Gaussian distributions, and some transformations of bivariate Gaussian models. These models are labeled $M_1 - M_{12}$. Here, we only provide some brief information on their notation and construction. Appendix B contains some further details and analytical formulas for the shape of the corresponding monotonic dependence functions.

$M_1:$ t-symm(ν,ρ); symmetric bivariate Student t distribution with ν degrees of freedom.

$M_2: LR1(\theta)$; linear regression of X on Y: $X = \theta Y + Z$, where Y and Z are independent $\mathcal{N}(0,1)$ variables and $\theta \leq 0$; *cf.* Zhu *et al.* [49].

M_j : $j = 3,4,5,$ and 8, $TBN(j,\rho)$; transformed bivariate normal distribution. Following Kowalczyk *et al.* [28], we consider $(X, f_j(Y))$, where (X,Y) has a bivariate normal distribution with parameter vector $(EX, EY, \mathrm{Var}X, \mathrm{Var}Y, \rho)$, while $f_3(y) = \mathrm{sgn}(y)[1-\exp\{-|y|\}]$, $f_4(y) = \mathrm{sgn}(y)[\exp\{|y|\}-1]$, and $f_5(y) = f_8(y) = \exp\{y\}$. In the models M_3, M_4 and M_5, the parameter vector was set to be $(0,0,1,1,\rho)$, while in the model M_8 it was $(0,0,4,4,\rho)$. Note that $TBN(5,\rho)$ and $TBN(8,\rho)$ are bivariate log-normal models, which only differ according to the variances of the marginal distributions.

$M_6:$ t-skew(ν,γ_1,γ_2); skew bivariate Student t distribution with ν degrees of freedom.

$M_7: Exp(\theta)$; bivariate exponential distribution with the distribution function
$F(x,y) = 1 - e^{-x} - e^{-y} + e^{-x-y-\theta xy}$, for $x, y > 0$.

$M_9: QR(\theta)$; quadratic regression of Y on X: $Y = \theta X^2 + Z$, where X and Z are independent, $X \sim U[-1,1]$, $Z \sim U\left[-\frac{1}{2},\frac{1}{2}\right]$, $\theta \in \left[0,\frac{3}{4}\right]$. Here, $U[a,b]$ denotes the uniform distribution on $[a,b]$.

$M_{10}:$ Mix(δ,ρ); mixture of three bivariate normal distributions P_{X_i,Y_i}, $i = 1,2,3$. This mixture is given by the formula $(1-2\eta)P_{X_1,Y_1} + \eta P_{X_2+\delta,Y_2-\delta} + \eta P_{X_3-\delta,Y_3+\delta}$, where $\eta = 0.05$ and $(X_1,Y_1),(X_2,Y_2),(X_3,Y_3)$ are independent bivariate normal random variables each with the parameter vector $(0,0,1,1,\rho)$; *cf.* Kowalczyk *et al.* [28].

$M_{11}: LR2(\theta)$; linear regression of X on Y: $X = \theta Y + Z$, Y and Z are independent, $Y \sim U[-1,1]$, $Z \sim U[-1,1]$, $\theta \in [-1,0]$; see Zhu *et al.* [49].

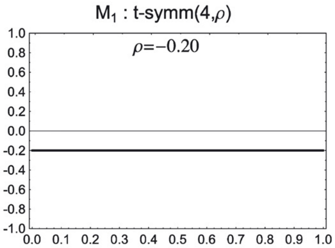
$M_{12}: U(\theta)$ uniform distribution on the set $A = \{(x,y) : x \in [-1,1],\ y \in [0,(1-x^2)^\theta]\}$, $\theta \geq 0$.

It is worth noticing that model M_2 is based on a bivariate normal distribution with parameter vector $(0,0,1,1,\rho(\theta))$, where $\rho(\theta) = \theta/\sqrt{\theta^2+1}$. The transformations of the bivariate normal models defining M_3 and M_4 correspond to Y having longer and shorter tails, respectively, compared to the normal distribution. In models M_5 and M_8, the right tails of the distribution of Y are made longer by the transformation. Model M_{12} supplements a series of illustrative cases describing the shape of $\mu_{X,Y}$ for uniform distributions on certain sets as presented in Kowalczyk *et al.* [28].

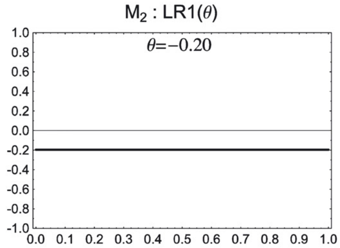
In Figure 1 we present the monotonic dependence functions for each of the models, along with the empirical significance levels and powers for a sample size $n = 200$. In the case of the test statistics $S^-(\epsilon)$ and $R^-(\epsilon)$, we give results corresponding to $\epsilon = 0.100$. In the case of the test statistic T, we proceed exactly as in Zhu *et al.* [49] and omit the 0.5% most extreme observations. We have also considered larger sample sizes and other trimming levels, *i.e.* different values of ϵ defining the test statistics $S^-(\epsilon)$ and $R^-(\epsilon)$. In Tables C.1–C.3 of Appendix C we present the empirical significance levels and powers for the following sample sizes: $n = 200, n = 400$ and $n = 1000$. In the cases $n = 400$ and $n = 1000$, the parameters of the alternatives $M_1 - M_{12}$ were chosen to guarantee that the test based on T has practically the same empirical power as it has when $n = 200$ with the parameter values given in Figure 1. The corresponding values of the parameters for larger sample sizes are given in Tables C.2–C.3. For each sample size, we considered values of the test statistics $R^-(\epsilon)$ and $S^-(\epsilon)$ corresponding to $\epsilon = 0.100,\ 0.075$ and 0.050. The codes were written in C Sharp.

Below, we briefly comment on some of the shapes of the monotonic dependence functions in Figure 1 and the relation between these shapes and the expected powers of the new tests. Further discussion on empirical significance levels and powers is postponed to Sections 4.2 and 4.3, respectively.

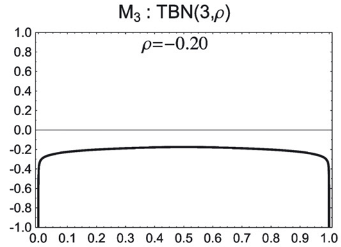
As we are now considering alternatives to EQD^+, it is useful to recall that EQD^- denotes negative expectation dependence, which means that $E(Y|X \leq x_p) \geq EY$ for all $p \in (0,1)$. Any EQD^- distribution corresponds



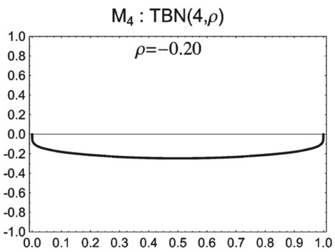
ρ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.044	0.052	0.048
-0.20	0.493	0.637	0.708



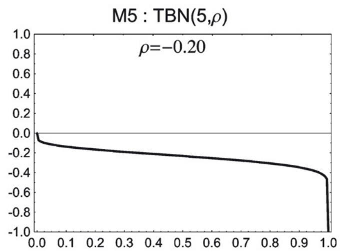
θ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.040	0.044	0.049
-0.20	0.749	0.793	0.791



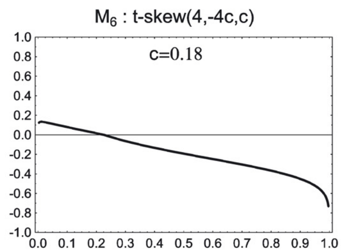
ρ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.053	0.048	0.055
-0.20	0.701	0.793	0.796



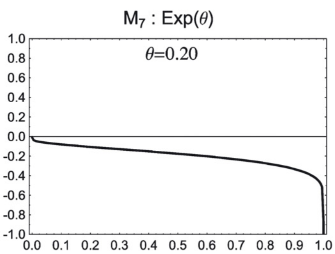
ρ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.042	0.046	0.046
-0.20	0.676	0.723	0.725



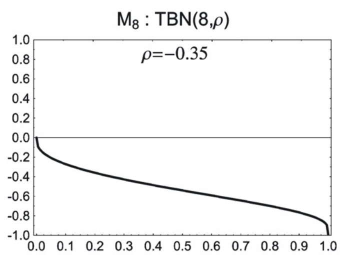
ρ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.076	0.041	0.038
-0.20	0.614	0.635	0.623



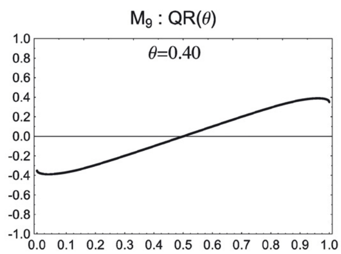
c	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.041	0.042	0.049
-0.18	0.529	0.687	0.707



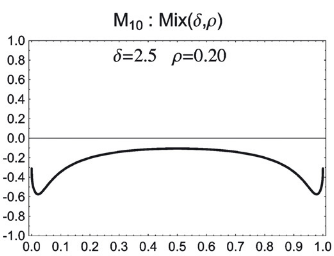
θ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.062	0.049	0.046
0.20	0.668	0.671	0.593



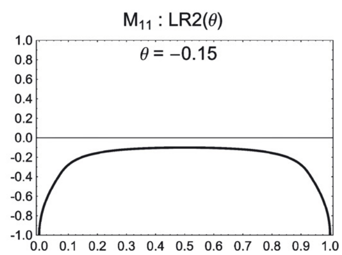
ρ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.076	0.034	0.020
-0.35	0.722	0.649	0.574



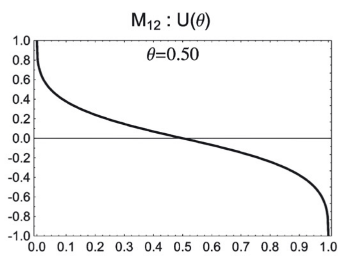
θ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.045	0.055	0.054
0.40	0.845	0.830	0.689



ρ/δ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0/0	0.055	0.058	0.060
0.20/2.5	0.811	0.752	0.644



θ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.058	0.058	0.054
-0.15	0.746	0.662	0.490



θ	$S^{-}(\epsilon)$	$R^{-}(\epsilon)$	T
0	0.060	0.057	0.059
0.50	0.906	0.780	0.452

FIGURE 1. Monotonic dependence functions for models $M_1 - M_{12}$ with the corresponding parameters displayed under each figure. The empirical significance levels (*middle row of the tables*) and empirical powers (*last row*) are given for $\epsilon = 0.100$ and $n = 200$; $\alpha = 0.05$.

to the alternative hypothesis, $\mathcal{H}_1$. Obviously, $\mathcal{H}_1$ covers all distributions for which EQD^+ is violated for at least one point $p \in (0,1)$. In simple terms, we should remember that, under $\mathcal{H}_1$, increasing the local strength of dependence at some point p means that $\mu_{Y,X}$ is negative at this point and increasing in absolute value.

For the distribution defined by M_1, the correlation coefficient is equal to ρ, while $E(Y|X) \doteq \rho X$. Therefore, from points 7 or 9 of Lemma 2.1, $\mu_{Y,X}(p) \equiv \rho$. The same properties apply to M_2. Note that Kowalczyk *et al.* [28] discuss $\mu_{Y,X}$ in the general case of the symmetric bivariate Student distribution, including the Cauchy distribution.

From property 9, we infer that models M_3-M_{12} have non-linear regression functions. For illustration, consider the last two models on the list. One can easily check that for M_{11} the conditional expectation $E(Y|X=x)$ is non-increasing and convex for $x < x_{0.5}$, while it is non-increasing and concave for $x > x_{0.5}$. For M_{12}, the conditional expectation $E(Y|X=x)$ is increasing and concave for $x < x_{0.5}$, while it is decreasing and concave for $x > x_{0.5}$. The signs and regions of monotonicity of $\mu_{Y,X}$ can be determined from point 9 of Lemma 2.1 and are reflected in the shape of the corresponding $\mu_{Y,X}$ in Figure 1.

Model M_3 illustrates how the local strength of EQD^- according to the bivariate Gaussian model M_2 changes when equal portions of the probability mass of the distribution of Y are transferred to each tail. Obviously, the strength of dependence should thus be larger for both small and large values of p and, indeed, we see such an effect. Analogously, when the mass of Y is only reallocated towards the right tail, in comparison to M_2, which takes place in the case of models M_5 and M_8, the strength of EQD^- dependence only increases for large values of p. In contrast, the transformation corresponding to M_4 shrinks the probability density of the distribution of Y based on M_2 towards the point 0 and, in effect, the measure of the strength of EQD^- decreases for values of p close to 0 or 1 and increases for values of p around 0.5.

In cases where the strength of EQD^- is large near at least one of the ends of (0,1), one might expect that the weighted variants of the Kolmogorov$-$Smirnov test for EQD^+ should lead to more powerful tests than the unweighted version of this statistic, T; *cf.* the comment on the relation between T and $L_{Y,X}(p)$ at the beginning of this section. Indeed, such a situation occurs for models $M_7 - M_{12}$. In the cases of M_3 and M_5, the interval where $\mu_{Y,X}$ is large is very narrow and so closely located to one or both end points of (0,1) that it cannot be detected well using either $S^-(0.100)$ or $R^-(0.100)$. Note also that heavier tails than Gaussian ones, asymmetric distributions, and uncorrelated but dependent random variables are common in many current applications. For more discussion, see Cont [8], Baur [4] and related papers. Therefore, the detection of alternatives of such types is of vital interest.

We close this Section with some remarks on possible choices of a weight for sup-type statistics. In many problems, a form of studentization is useful. For example, when testing for uniformity, this idea leads to the simple weight function $\sqrt{p(1-p)},\ p \in (0,1)$, *i.e.* an Anderson–Darling sup-type statistic. The impact of this weight on the power of the resulting test is easily predictable, at least qualitatively. In our problem, the form of the asymptotic variance of $L_{Y,X}(p)$ is very complex (*cf.* the comment following Prop. 3.2) and there is little chance that some plug-in estimator results in a sensitive and stable weighted test statistic. Besides, the following question is even more important: what kind of deviations from the null hypothesis might be detected more reliably after applying such a weight function. A big advantage of applying $M_Y^-(p)$, is the predictable effect of the weight on behavior of the appropriate test, at least in some important cases that we have discussed above. Obviously, these effects depend on the form of the monotonic dependence function.

As mentioned in Section 4.1, other weights can be considered as well, since in the proof of the validity of the resampling method only the weak law of large numbers is applied to the denominator defining the respective analogue of $R^-(\epsilon)$. However, we do not wish to overload this paper with further technicalities and simulations. Our contribution casts new light into some aspects of testing for EQD^+, while our tests provide some improvement in comparison to the existing one, as shown in Section 4.2. Our approach also opens space to develop new tests. For example, one could consider the integral of the negative part of $\hat{\mu}_{Y,X}(p)$ over (0,1), which would result in an integral-type Anderson–Darling statistic for EQD^+. However, this direction of research requires elaborating new ways of proving the validity of the resampling method and should be the subject of future work.

4.2. Empirical significance levels

Theorem 3.4 shows that to control the empirical significance level of our tests, it is sufficient to carry out simulations based on the models M_1 - M_{12} with parameters which ensure that $\mu_{Y,X}(p) \equiv 0$. Their parametrizations correspond to the appropriate parameters being simply equal to zero. These parameters are displayed in the first row of the tables in Figure 1. Following Scaillet (2005), we call a fraction of rejections of $\mathcal{H}_0$ under $\mu_{Y,X}(p) \equiv 0$ the empirical significance level. The Monte Carlo experiments show that the multiplier method yields appropriate empirical significance levels in the majority of the cases considered. Some deficiencies are observed for $S^-(\epsilon)$ under long-tailed alternatives (M_1, M_5, M_6 and M_8) when $n = 200$. For larger sample sizes the situation improves, except in the case M_8. In this case, T also yields inappropriate empirical significance levels for all the sample sizes under consideration; *cf.* Tables C.1−C.3. We have also applied a classical bootstrap algorithm in a preliminary study, but we abandoned it, since the results of the multiplier method were more encouraging.

4.3. Empirical powers

In Figure 2, we present a graphical summary of the empirical powers of the tests based on T, $S^-(0.100)$ and $R^-(\epsilon)$ with $\epsilon = 0.050, 0.750, 0.100$ for the three cases $n = 200, n = 400$ and $n = 1000$ when the nominal significance level is 5%. The vertical bars represent the differences between the powers of the test based on T and the powers of the new tests for the corresponding ϵ's and n's.

These outcomes show that in most cases the new test based on $R^-(\epsilon)$ with $\epsilon = 0.050, 0.750$ and 0.100 is comparable to or better than the existing test based on T. The only exception is for the model based on the symmetric bivariate Student distribution with four degrees of freedom. Although the monotonic dependence function for the Student distribution is almost the same as for the bivariate normal model M_2, the heavier tails of the marginals influences the magnitude of the empirical processes defining (3.4) and make it harder to reject the null hypothesis. However, this deficiency disappears relatively quickly as the sample size increases. On the other hand, the noticeable deficiencies of the test based on T for models $M_7 - M_{12}$ only decrease marginally as n increases. The test based on $S^-(\epsilon)$ is less stable than the one based on $R^-(\epsilon)$ in the sense that it gives greater gains in some cases, but the loss of power (with respect to the test based on T) in some other situations are also greater. As n increases the comparative empirical power of the test based on $S^-(\epsilon)$ stabilizes.

To obtain a more concise picture of the empirical powers of these three tests as indicated by Figure 2, we calculated the average powers of $S^-(0.100), R^-(0.100)$ and T over the twelve alternatives for each sample size n. The experiments summarized in Figure 2 also indicate that it is beneficial to slowly decrease ϵ as n increases. Our empirical observation is that $\epsilon = \epsilon(n) = \sqrt{2/n}$ is a reasonable choice. Therefore, we additionally considered the statistic $R^-(\epsilon(n))$ in this comparison. For each of the three weighted statistics, we calculated the difference between the average power of the corresponding test and the test based on T. These results are presented in Figure 3.

This short summary clearly shows that introducing weighting in any of the three forms considered in Figure 3 is beneficial. Moreover, the variant $R^-(0.100)$ is more powerful (on average) than $S^-(0.100)$ and can thus be recommended as a well studied improvement to the test based on T. Note also that the simulation results presented in Tables C.1−C.3 show that the empirical significance levels for the tests based on $R^-(\epsilon(n))$ are stable and satisfactorily precise, while the average powers of the tests based on $R^-(\epsilon(n))$ are the greatest, see Figure 3, and show the appealing tendency of growing with n. Therefore, although our theoretical considerations do not cover this case, the test based on $R^-(\epsilon(n))$ may also be considered to be a useful solution.

5. Example based on real data

We shall consider the Danish fire insurance data set available at http://www.ma.hw.ac.uk/~mcneil/data.html. These data consist of losses in Danish Krone for the years 1980 to 1990. This is a multivariate data set containing the financial loss to buildings X_i, to contents Y_i and to profits Z_i caused by fire number i,

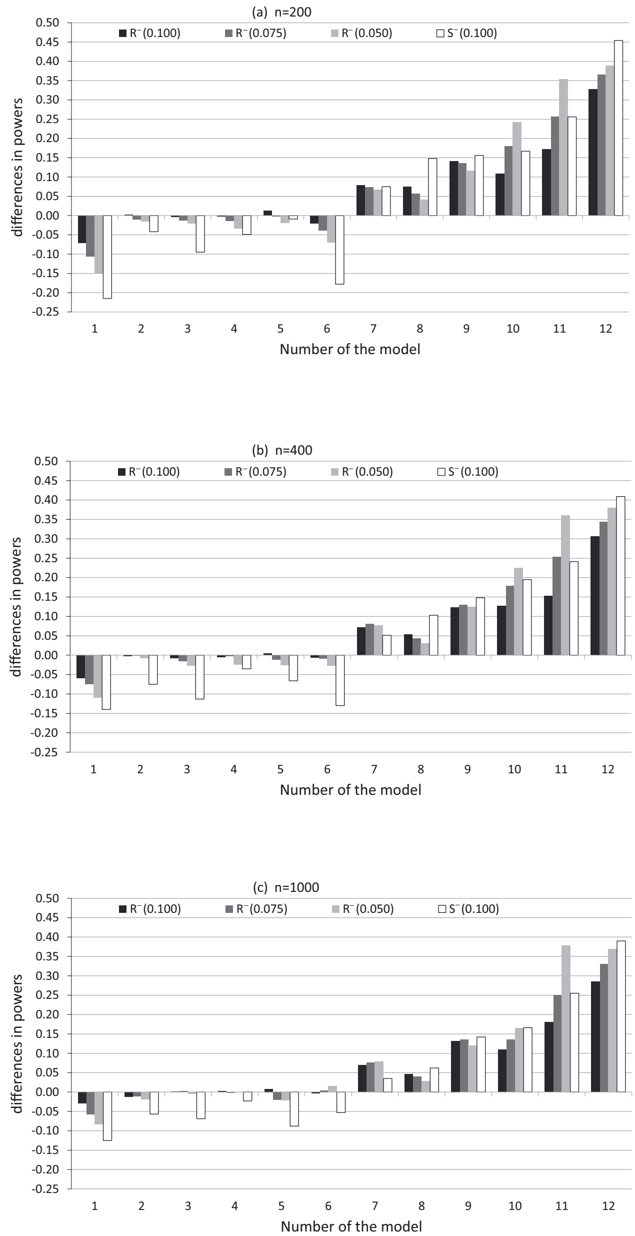


FIGURE 2. Differences between the empirical powers of the tests based on $S^{-}(0.100)$, $R^{-}(\epsilon)$ and T for given values of ϵ and n; $\alpha = 0.05$, alternatives $M_1 - M_{12}$ with adjusted parameters under $n = 400$ (b), and $n = 1000$ (c).

$i = 1, \ldots, 1502$. A positive dependence between these losses is expected, since the insured objects were exposed to the same cause of damage. Therefore, the hypothesis of positive quadrant dependence (PQD) for these data was formulated and verified in Scaillet [41], Gijbels *et al.* [20], and Ledwina and Wyłupek [34]. We concentrate below mainly on inference for the losses $(X_i, Y_i),\ i = 1, \ldots, 1502$. Though some classical tests do not provide any evidence against either the independence or PQD of these variables, the most specialized tests for PQD reject it with p-values equal to 0. One explanation of these conflicting conclusions is that the dependencies among the

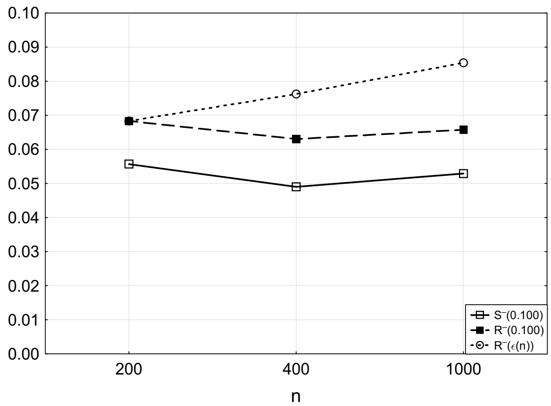


FIGURE 3. Average gains in the power of the tests based on $S^-(0.100)$, $R^-(0.100)$ and $R^-(\epsilon(n))$ in comparison to the test based on T according to n; $\alpha = 0.05$.

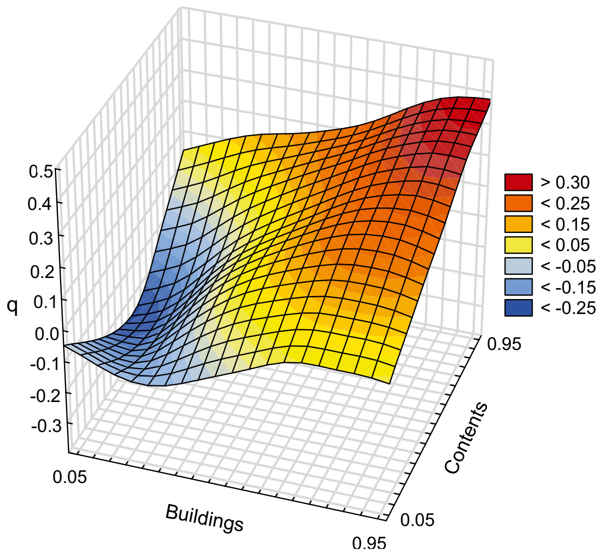


FIGURE 4. Danish fire insurance data; $n = 1502$. Smoothed estimator of the dependence measure q for loss to buildings X and loss to content Y.

variables are more complex than some approximately linear trend and had not been detected by the relatively simple tools, that had been previously applied. To gain more insight into the dependence structure of these data see Figure 4. This illustrates an estimator of the function valued measure of dependence q proposed in Ledwina [33]. This measure is defined on $(0,1) \times (0,1)$ to be the appropriately standardized difference between the underlying and independence copulas. This measure is non-negative (non-positive) if and only if PQD (NQD – negative quadrant dependence) exists, while independence is equivalent to $q \equiv 0$. Several useful properties of this measure are presented in Ledwina [33]. Here, we would like to add that the measure q is equivalent to applying the concept of bivariate quantilogram (see Linton and Whang, [36]) to a bivariate copula. In view of the evidence in Figure 4, it is clear that it is unreasonable to infer either independence or PQD for these data. For a detailed analysis of the hypothesis of PQD for these data, see Ledwina and Wyłupek [34].

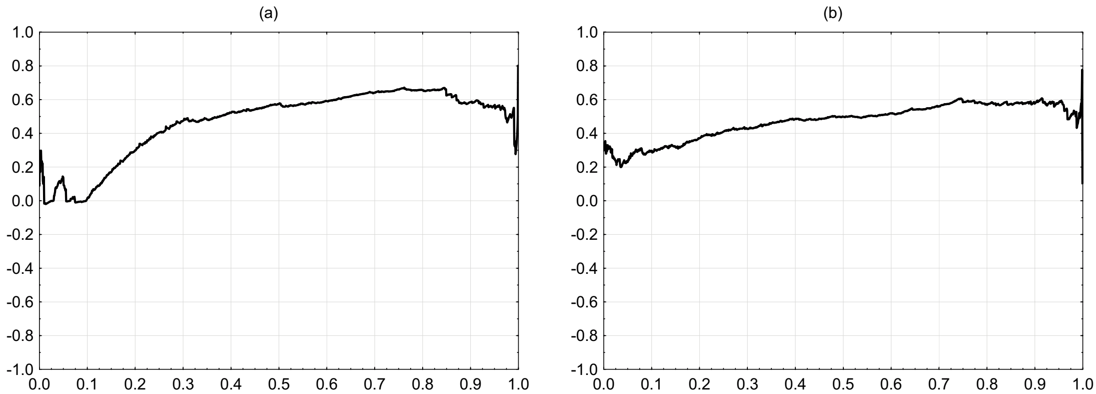


FIGURE 5. Danish fire insurance data; $n = 1502$. Empirical monotonic dependence functions: (a) Y on X; (b) X on Y.

TABLE 1. Danish fire insurance data; $n = 1502$. Estimated p-values from testing EQD^+ of Y on X and EQD^+ of X on Y based on $S^-(\epsilon), R^-(\epsilon)$ and T, $\epsilon = 0.050, 0.075, 0.100$ and $\epsilon = \epsilon(n) = 0.036$,. Based on 10 000 Monte Carlo runs.

Test	S^-(0,100)	S^-(0,075)	S^-(0,050)	S^-(0,036)	R^-(0,100)	R^-(0,075)	R^-(0,050)	R^-(0,036)	T
p-value for EQD^+ of Y on X	0.7598	0.7027	0.7375	0.7487	0.7603	0.6998	0.7330	0.7439	0.8131
p-value for EQD^+ of X on Y	1	1	0.9998	0.9980	1	1	0.9997	0.9970	0.9974

In view of the above evidence, it is interesting to see whether or not a weaker form of positive dependence than PQD, specifically positive expectation dependence EQD^+, is appropriate for these data. In Figure 5 we present estimates of the monotonic dependence functions $\mu_{Y,X}(p)$ and $\mu_{X,Y}(p)$, $p \in [0.001, 0.999]$, for the $n = 1502$ positive claims. Both curves are non-negative for all $p \geq 0.05$ and indicate a non-decreasing trend for $p \in (0.1, 0.8)$. It can also be seen that the estimated dependence is not symmetrical with respect to X and Y. When X is relatively small, the curves suggest that the dependence of Y on X is stronger than the dependence of X on Y for claims in a similar range. On the other hand, the dependence of X on Y is slightly stronger that the dependence of Y on X for moderate and large values of Y.

In Table 1 we present conclusions from tests of the hypotheses $\mu_{Y,X}(p) \geq 0$ and $\mu_{X,Y}(p) \geq 0$. These tests provide no evidence against the postulated type of positive dependence, while the p-values indicate very good fit.

Similar analysis was also performed for the other two pairs of variables that may be constructed from the trio of variables observed. In the case of losses to content and losses to profit, the appropriate hypotheses of PQD, and thus EQD^+, were accepted with the corresponding p-values all being close to 1. In the case of losses to buildings and losses to profit, the hypothesis of PQD is rejected at standard levels of significance, while the weaker hypothesis of EQD^+ is again accepted, based on p-values being practically equal to 1 for the hypothesis of EQD^+ for loss to profit on loss to buildings and at least 0.8433 for EQD^+ for loss to buildings on loss to profit.

6. Summary and concluding remarks

In this paper, we have proposed nonparametric tests for positive expectation dependence. Our contributions include the following: First, we proposed a useful reparametrization of the test and then inference based on a monotonic dependence function that measures the local strength of dependence. This step helps us to propose an appropriate function for weighting the test statistic in a natural Kolmogorov−Smirnov type test and proves that testing can be accompanied by graphical evidence of departures from the null hypothesis. Second, we showed that the multiplier method gives sufficiently precise control of the significance level for finite samples. Our solutions were compared to the currently used test procedure in an extensive simulation study and proved to be competitive.

By introducing such a reparametrization, we introduce nuisance parameters into the analysis, which make the argument more complex. These nuisance parameters are the pth quantiles, x_p and y_p, $p \in (0,1)$, of the unknown marginal distributions. In our test procedure, we restrict attention to some subinterval of $(0,1)$. In the case where one is interested in EQD^+ dependence based on the concept of a copula or other bivariate distribution with known marginals, similar analysis is possible and expected to be less involved.

Other solutions, such as integral-type statistics for the reparametrized problem, seem to be of interest for further study. Also, it would be useful to develop nonparametric confidence bounds for the monotonic dependence function. This would formalize the proposed graphical analysis of possible sources of deviation from positive expectation dependence. Extensions of the approach to dependent samples would also be welcome.

Appendix A. Mathematical proofs

A.1. Proof of Proposition 3.2

First, consider the process Z_n and observe that

$$\sup_{\epsilon\le p\le 1-\epsilon}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^{n} Y_i\left[(np-\lfloor np\rfloor)I(X_i=\hat{x}_p)\right]\right| \le \max_{1\le i\le n}\left\{\frac{1}{\sqrt{n}}|Y_i|\right\}.$$

The assumption $EY^2<\infty$ implies that $\max_{1\le i\le n}\{|Y_i|/\sqrt{n}\}=o_P(1)$. Thus we can restrict attention to

$$\lambda_n(p)=\frac{1}{n}\left[p\sum_{i=1}^{n} Y_i-\sum_{i=1}^{n} Y_i I(X_i<\hat{x}_p)\right]$$

and investigate the weak convergence of $\sqrt{n}(\lambda_n-l_{Y,X})$. For this purpose, recall that $X_{(1)}<\ldots<X_{(n)}$ are the order statistics for the sample $X_1,\ldots,X_n$ and denote the induced order statistics in the sample $Y_1,\ldots,Y_n$ by $Y_{[1]},\ldots,Y_{[n]}$. Thus

$$\lambda_n(p)=\frac{1}{n}\left(p\sum_{i=1}^{n} Y_{[i]}-\sum_{i=1}^{\lfloor np\rfloor} Y_{[i]}\right).$$

Set

$$\xi_n(p)=\frac{1}{n}\sum_{i=1}^{\lfloor np\rfloor} Y_{[i]}.$$

From Theorem 1 of Davydov and Egorov [11], restricted to $[\epsilon,1-\epsilon]$, it holds that $\sqrt{n}(\xi_n-f)\Rightarrow\bar{\xi}$ in $D[\epsilon,1-\epsilon]$, where $f(p)=\int_0^p m(s)\mathrm{d}s=\int_0^p E(Y|F(X)=s)\mathrm{d}s$ and $\bar{\xi}$ is a centered Gaussian process concentrated on $C[\epsilon,1-\epsilon]$. The process $\bar{\xi}$ has the structure given by (5) and (7) of Theorem 1, while its covariance function is presented on pp. 302−303 ibidem. Since $\lambda_n(p)=[p\xi_n(1)-\xi_n(p)]$, the continuous mapping theorem and the above imply that $Z_n\Rightarrow\bar{\xi}$ in $D[\epsilon,1-\epsilon]$. However, the process Z_n takes values in $C[0,1]$ and therefore from Billingsley [7], page 39, we obtain the first statement of Proposition 3.2.

The convergence

$$\sup_{\epsilon\le p\le 1-\epsilon}\left[M_Y^+(p)-m_Y^+(p)\right]\to 0 \quad \text{in probability}$$

follows from Lemma 3 in Bednarski and Ledwina [5]. The same conclusion holds for M_Y^- and m_Y^-. The rest of the proof is an immediate consequence of the properties of weak convergence. □

A.2. Proof of Proposition 3.3

Proof of Part (i): From the definitions of $S^-(\epsilon)$ and $-V_n^-(p)$, we obtain

$$S^-(\epsilon) \leq \sup_{\epsilon \leq p \leq 1-\epsilon} \{-V_n^-(p)\} + \sup_{\epsilon \leq p \leq 1-\epsilon} \left\{-\sqrt{n}\frac{l_{Y,X}(p)}{M_Y^-(p)}\right\}.$$

Under $\mathcal{H}_0$ it holds that $l_{Y,X}(p) \geq 0$ for all p. Hence, $S^-(\epsilon) \leq \sup_{\epsilon \leq p \leq 1-\epsilon}\{-V_n^-(p)\}$ and equality holds if $l_{Y,X}(p) \equiv 0$ on $[\epsilon, 1-\epsilon]$. Therefore, Proposition 3.2 and the continuous mapping theorem yield (3.2).

Proof of Part (ii): For $p_0 \in [\epsilon, 1-\epsilon]$, assuming that $\mathcal{H}_0$ does not hold, we have $\Pr(S^-(\epsilon) > c) \geq \Pr(-\sqrt{n}\hat{\mu}_{Y,X}^-(p_0) > c)$. Since $c > 0$ and $\mu_{Y,X}(p_0) < 0$, application of Proposition 3.2 concludes the proof.

A.3. Proof of Theorem 3.4

The analysis of the asymptotics of $\hat{p}(n, \epsilon)$ under $\mathcal{H}_0$ is carried out *via* the three steps elaborated in Appendices A.3.1−A.3.3. The main idea of the proof is to reduce considerations to the case in which the theoretical quantiles of X are known and then apply the elegant result of Zinn on the almost sure central limit theorem for variables taking values in a separable Banach space, published in Ledoux and Talagrand [31,32]. Appendix A.3.4 justifies the hypothesis regarding the asymptotic behavior of $\hat{p}(n, \epsilon)$ under $\mathcal{H}_1$.

We start with an analysis of the numerator of $S^{U-}(p)$.

A.3.1. Approximation of $L_{Y,X}^U(p)$

The approach is asymptotic and conditional on the realization of the sample $(X_1, Y_1), \ldots (X_n, Y_n)$.

Set $Y_i^0 = Y_i - EY_i$ and write

$$\begin{aligned}\sqrt{n}L_{Y,X}^U(p) &= \lambda_{1n}(p) + \frac{1}{\sqrt{n}}\sum_{i=1}^n U_i\Big\{l_{Y,X}(p) - L_{Y,X}(p)\Big\} \\ &\quad + \frac{1}{\sqrt{n}}\sum_{i=1}^n U_i(\bar{Y} - EY)\Big[p - I(X_i < \hat{x}_p) - (np - \lfloor np \rfloor)I(X_i = \hat{x}_p)\Big], \end{aligned} \tag{A.1}$$

where

$$\lambda_{1n}(p) = \frac{1}{\sqrt{n}}\sum_{i=1}^n U_i\Big\{Y_i^0\Big[p - I(X_i < \hat{x}_p) - (np - \lfloor np \rfloor)I(X_i = \hat{x}_p)\Big] - l_{Y,X}(p)\Big\}.$$

For any $\delta \in (0, \infty)$

$$\Pr^U\left(\sup_{\epsilon \leq p \leq 1-\epsilon}\left|\frac{1}{\sqrt{n}}\sum_{i=1}^n U_i\left\{l_{Y,X}(p) - L_{Y,X}(p)\right\}\right| > \delta\right) \leq \left[\sup_{\epsilon \leq p \leq 1-\epsilon}|l_{Y,X}(p) - L_{Y,X}(p)|\right]^2 \times \left[\delta^{-2}E\left(\frac{1}{n}\sum_{i=1}^n U_i^2\right)\right]. \tag{A.2}$$

From Proposition 3.2, the first factor in (A.2) tends to 0 for almost all sequences $(X_1, Y_1), (X_2, Y_2), \ldots$ Since $EU_i^2 = 1$, $i = 1, \ldots, n$, the second component of (A.1) is negligible conditional on almost any sequence. A similar conclusion holds for the third component of (A.1).

Now observe that

$$\lambda_{1n}(p) = \lambda_{2n}(p) + \frac{1}{\sqrt{n}}\sum_{i=1}^n U_iY_i^0\Big[p - F(\hat{x}_p)\Big] + \frac{1}{\sqrt{n}}\sum_{i=1}^n U_i\Big[l_{Y,X}\big(F(\hat{x}_p)\big) - l_{Y,X}(p)\Big], \tag{A.3}$$

where

$$\lambda_{2n}(p) = \frac{1}{\sqrt{n}}\sum_{i=1}^n U_i\Big\{Y_i^0\Big[F(\hat{x}_p) - I(X_i < \hat{x}_p) - (np - \lfloor np \rfloor)I(X_i = \hat{x}_p)\Big] - l_{Y,X}\big(F(\hat{x}_p)\big)\Big\}.$$

Since F and $l_{Y,X}$ are continuous functions and $\sup_{\epsilon\le p\le 1-\epsilon}\big|\hat{x}_p - x_p\big| \to 0$ almost surely, arguing as above, we conclude that the second and third components of (A.3) are negligible.

Finally, note that

$$\lambda_{2n}(p) = \lambda_{3n}(p) - \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i\Big\{Y_i^0\Big[np - nF(\hat{x}_p) + \lfloor nF\big(\hat{x}_p\big)\rfloor - \lfloor np\rfloor\Big]I(X_i = \hat{x}_p)\Big\} + o_P(1), \tag{A.4}$$

where

$$\lambda_{3n}(p) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i\Big\{Y_i^0\Big[F(\hat{x}_p) - I\big(F(X_i) < F(\hat{x}_p)\big)\Big] - l_{Y,X}\big(F(\hat{x}_p)\big)\Big\}.$$

The remainder in (A.4) is stochastically small, while the order of the error term, $o_P(1)$, follows from a similar argument to the one given in Section A.1.

Summarizing (A.1)–(A.4), we conclude that

$$\sqrt{n}L_{Y,X}^U(p) = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} U_i\Big\{Y_i^0\Big[F(\hat{x}_p) - I\big(F(X_i) < F(\hat{x}_p)\big)\Big] - l_{Y,X}\big(F(\hat{x}_p)\big)\Big\} + r_n, \tag{A.5}$$

where $r_n = o_P(1)$, conditional on almost any sample.

A.3.2. Approximation of the distribution of $S^{U-}(\epsilon)$

Recall that

$$S^{U-}(\epsilon) = \sup_{\epsilon\le p\le 1-\epsilon}\left\{-\sqrt{n}\frac{L_{Y,X}^U(p)}{M_Y^-(p)}\right\}. \tag{A.6}$$

The aim of this section is to show that, conditional on the sample $(X_1,Y_1),\dots,(X_n,Y_n)$, the asymptotic behavior of $S^{U-}(\epsilon)$ under $\Pr^U$ is the same as that of the supremum of an appropriately defined variant of $V_n^-(p)$, constructed on the basis of the sample $(X_1,Y_1),\dots,(X_n,Y_n)$, *cf.* (A.11) below. A crucial step in obtaining this result is to approximate the supremum in (A.6) by the supremum of a corresponding process with values in $C[\epsilon, 1-\epsilon]$ and to reduce the problem to a case in which the theoretical quantiles are known.

Observe first that the process described in (A.5) depends on p *via* the piecewise constant function $\hat{x}_p$. Moreover, from Proposition 3.2, $M_Y^- \to^{\Pr} m_Y^-$. Also note that, in any derivation of the asymptotics, we have the freedom to change the denominator in (A.6), as long as the limit m_Y^- is retained. Moreover, when the analysis solely concerns randomness due to $U_1,\dots,U_n$, we may consider a deterministic denominator. In addition, for large n, without loss of generality, we can approximate the continuous function m_Y^- by a continuous, piecewise linear function on the intervals $(0, F(X_{(1)}), [F(X_{(k)}), F(X_{(k+1)})],\ k = 1,\dots,n-1, (F(X_{(n)}), 1)$. Let us denote such a function by $\tilde{m}_Y^-(p)$. The result (A.5) and the above argument show that instead of analyzing $S^{U-}(\epsilon)$, it suffices to analyze

$$S_3^{U-}(\epsilon) = \max_{t\in\mathcal{T}} \frac{\frac{-1}{\sqrt{n}}\sum_{i=1}^{n} U_i\Big\{Y_i^0\Big[t - I\big(F(X_i) < t\big)\Big] - l_{Y,X}(t)\Big\}}{\tilde{m}_Y^-(t)},$$

where $\mathcal{T} = \big\{t = F(X_{(\lfloor\epsilon n\rfloor+1)}), F(X_{(\lfloor\epsilon n\rfloor+2)}),\dots,F(X_{(\lfloor(1-\epsilon)n\rfloor+1)})\big\}$.

Consider the following two auxiliary processes

$$\upsilon_n^U(t) = \frac{1}{\sqrt{n}\tilde{m}_Y^-(t)}\sum_{i=1}^{n} U_i\Big\{Y_i^0\Big[t - I\big(F(X_i) < t\big)\Big] - l_{Y,X}(t)\Big\}, \tag{A.7}$$

and

$$\upsilon_n(t) = \frac{1}{\sqrt{n}\tilde{m}_Y^-(t)}\sum_{i=1}^{n}\Big\{Y_i^0\Big[t - I\big(F(X_i) < t\big)\Big] - l_{Y,X}(t)\Big\}, \tag{A.8}$$

together with their polygonal interpolations

$$\omega_n^U(t) = \frac{1}{\sqrt{n}\tilde{m}_Y^-(t)} \sum_{i=1}^n U_i \Big\{ Y_i^0 \Big[t - I\big(F(X_i) < t\big) + a_i(t) \Big] - l_{Y,X}(t) \Big\}$$

and

$$\omega_n(t) = \frac{1}{\sqrt{n}\tilde{m}_Y^-(t)} \sum_{i=1}^n \Big\{ Y_i^0 \Big[t - I\big(F(X_i) < t\big) + a_i(t) \Big] - l_{Y,X}(t) \Big\},$$

where $a_i(t) = \{1 + n^3[F(X_i) - t]\} \times I(F(X_i) < t \leq F(X_i) + n^{-3})$. It can be seen that $0 \leq a_i(t) \leq 1$ and

$$\sup_{t\in(0,1)} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i^0 a_i(t) \right| \to 0, \quad \sup_{t\in(0,1)} \left| \frac{1}{\sqrt{n}} \sum_{i=1}^n Y_i^0 U_i a_i(t) \right| \to 0$$

in probability, where the final relation concerns the probability measure $\Pr^U$ and holds for almost all $(X_1, Y_1), (X_2, Y_2), \ldots$. Using this notation, it follows that $S_3^{U-}(\epsilon) = \max_{t\in\mathcal{T}}\{-\upsilon_n^U(t)\}$. On the other hand, by assumption, $E(Y_i^0)^2 < \infty$. Therefore, from Theorem 10.14 of Ledoux and Talagrand [32], ω_n^U satisfies the central limit theorem in $C[\epsilon, 1-\epsilon]$, conditional on almost any sample $(X_1, Y_1), (X_2, Y_2), \ldots$, if and only if ω_n does. More precisely, $\omega_n \Rightarrow \omega$, if and only if $\omega_n^U \Rightarrow \omega^U$, conditional on the sample, where ω^U is the Gaussian process with the same covariance structure as ω. In addition, from part 1) of the proof of Theorem 1 of Davydov and Egorov [11], the above, and the continuous mapping theorem, it follows that $\omega_n \Rightarrow \omega$, in $C[\epsilon, 1-\epsilon]$, where ω is the Gaussian process with $E\omega(t) = 0$ and the covariance function is given on page 302 ibidem.

The above allow us to state that for any $c \in \mathbb{R}$

$$\left| \Pr^U \left(\max_{t\in\mathcal{T}}\{-\omega_n^U(t)\} > c \right) - \Pr \left(\max_{t\in\mathcal{T}}\{-\omega_n(t)\} > c \right) \right| \to^{\Pr} 0 \tag{A.9}$$

and

$$\left| \Pr^U \left(\max_{t\in\mathcal{T}}\{-\upsilon_n^U(t)\} > c \right) - \Pr \left(\max_{t\in\mathcal{T}}\{-\upsilon_n(t)\} > c \right) \right| \to^{\Pr} 0.$$

Observe also that

$$\max_{t\in\mathcal{T}}\{-\upsilon_n^U(t)\} = \sup_{\epsilon\leq p\leq 1-\epsilon} \left\{ -\frac{\lambda_{3n}(p)}{m_Y^-(F(\hat{x}_p))} \right\}, \tag{A.10}$$

$$\max_{t\in\mathcal{T}}\{-\upsilon_n(t)\} = \sup_{\epsilon\leq p\leq 1-\epsilon} \left\{ \frac{\frac{-1}{\sqrt{n}} \sum_{i=1}^n \Big\{ Y_i^0 \Big[F(\hat{x}_p) - I\big(F(X_i) < F(\hat{x}_p)\big) \Big] - l_{Y,X}\big(F(\hat{x}_p)\big) \Big\}}{m_Y^-(F(\hat{x}_p))} \right\}.$$

Since $m_Y^-(F(\hat{x}_p)/M_Y^-(p) \to 1$ in probability, the conclusion of this section is that

$$\left| \Pr^U \left(\sup_{\epsilon\leq p\leq 1-\epsilon} \left\{ -\sqrt{n} \frac{L_{Y,X}^U(p)}{M_Y^-(p)} \right\} > c \right) - \Pr \left(\sup_{\epsilon\leq p\leq 1-\epsilon} \left\{ -\sqrt{n} \frac{L_{Y,X}(p) - l_{Y,X}(F(\hat{x}_p))}{M_Y^-(p)} \right\} > c \right) \right| \to^{\Pr} 0. \tag{A.11}$$

A.3.3. Asymptotic analysis of $\hat{p}(n,\epsilon)$ under $\mathcal{H}_0$

From the definitions of the appropriate statistics

$$\hat{p}(n,\epsilon) = \Pr^U \left(S^{U-}(\epsilon) > S^-(\epsilon) \right) = \Pr^U \left(\sup_{\epsilon\leq p\leq 1-\epsilon} \left\{ -\sqrt{n} \frac{L_{Y,X}^U(p)}{M_Y^-(p)} \right\} > \sup_{\epsilon\leq p\leq 1-\epsilon} \left\{ -\sqrt{n} \frac{L_{Y,X}(p)}{M_Y^-(p)} \right\} \right). \tag{A.12}$$

Now recall that $l_{Y,X}(p) \geq 0$ under $\mathcal{H}_0$ and the same holds for $l_{Y,X}(F(\hat{x}_p))$ in the case where the sample $X_1, X_2, \ldots$ is fixed. Hence, (A.12) yields

$$\hat{p}(n,\epsilon) \geq \Pr^U\left(\sup_{\epsilon\leq p\leq 1-\epsilon}\left\{-\sqrt{n}\frac{L^U_{Y,X}(p)}{M^-_Y(p)}\right\} > \sup_{\epsilon\leq p\leq 1-\epsilon}\left\{-\sqrt{n}\frac{L_{Y,X}(p)-l_{Y,X}(F(\hat{x}_p))}{M^-_Y(p)}\right\}\right) \tag{A.13}$$

with equality holding if $l_{Y,X}(p) \equiv 0$ on $[\epsilon, 1-\epsilon]$. From (A.13) and (A.6)

$$\Pr\left(\hat{p}(n,\epsilon) < \alpha\right) \leq \Pr\left(\Pr^U\left(S^{U-}(\epsilon) > \sup_{\epsilon\leq p\leq 1-\epsilon}\left\{-\sqrt{n}\frac{L_{Y,X}(p)-l_{Y,X}(F(\hat{x}_p))}{M^-_Y(p)}\right\}\right) < \alpha\right). \tag{A.14}$$

Set

$$Z_{1n} = S^{U-}(\epsilon) \quad \text{and} \quad Z_{2n} = \sup_{\epsilon\leq p\leq 1-\epsilon}\left\{-\sqrt{n}\frac{L_{Y,X}(p)-l_{Y,X}(F(\hat{x}_p))}{M^-_Y(p)}\right\}.$$

The expression on the right hand side of (A.14) can be written as

$$\Pr\left(\Pr^U\left(Z_{1n} > Z_{2n}|Z_{2n}\right) < \alpha\right). \tag{A.15}$$

For $c \in \mathbb{R}$, set $D_n(c) = \Pr(\Pr^U(S^{U-}(\epsilon) < c))$. Using this notation, from Section A.3.2, it follows that $\lim\limits_{n\to\infty} D_n(c) = \lim\limits_{n\to\infty}\Pr(\Pr^U(\max_{t\in\mathcal{T}}\{-\omega_n^U(t)\} < c)) = \lim\limits_{n\to\infty}\Pr(\Pr^U(Z_{1n} < c))$. Hence, from (A.9) and (A.10), we infer that the limiting distribution functions of Z_{1n} and Z_{2n} coincide. In addition, the above yields

$$\lim_{n\to\infty}\Pr\Big(\hat{p}(n,\epsilon) < \alpha\Big) \leq \lim_{n\to\infty}\Pr\left(1 - D_n(Z_{2n}) < \alpha\right). \tag{A.16}$$

Observe that the process ω_n^U takes values in $C[\epsilon, 1-\epsilon]$, has mean 0 and is Gaussian. Its weak limit ω has the same properties. Therefore, from Theorem 1 of Tsirel'son [44], the distribution of $\max_{t\in\mathcal{T}}\{-\omega_n^U(t)\}$ is absolutely continuous on $(0,\infty)$. A similar conclusion holds for the weak limit of $\sup_{\epsilon\leq t\leq 1-\epsilon}\{-\omega(t)\}$. This ensures that $D_n(Z_{2n})$ satisfies $\lim_{n\to\infty}\Pr(1 - D_n(Z_{2n}) < \alpha) = \alpha$ for $\alpha < 1/2$. This proves (i) of Theorem 3.4.

A.3.4. Asymptotic analysis of $\hat{p}(n,\epsilon)$ under $\mathcal{H}_1$

From (ii) of Proposition 3.3 we infer that $S^-(\epsilon)$ tends to $+\infty$. On the other hand, $S^{U-}(\epsilon)$ is the supremum of a process with mean 0 and continuous sample paths. Therefore, it is finite with $\Pr^U$-probability 1. Hence, $\Pr(\hat{p}(n,\epsilon) < \alpha) = \Pr(\Pr^U(S^{U-}(\epsilon) > S^-(\epsilon)) < \alpha) \to 1$, as $n \to \infty$, and the proof is concluded.

Appendix B. Analytical formulas for $\mu_{Y,X}(p)$ for the models $M_1 - M_{12}$

First, we introduce some auxiliary notation. Let Φ and Φ^{-1} denote the $\mathcal{N}(0,1)$ distribution function and its inverse, respectively. Also, let $B_{(a,b)}$ and $B^{-1}_{(a,b)}$ denote the distribution function of the beta distribution with parameters a and b and its inverse, respectively.

The analytical forms of the monotonic dependence functions for some of the models we consider here are already known. In particular, Kowalczyk *et al.* [28] give the appropriate results for the transformed bivariate models M_5 and M_8, as well as the mixture model M_{10}. Since M_2 assumes a bivariate normal distribution, the formula for $\mu_{Y,X}(p)$ is a consequence of Theorem 2 in Kowalczyk and Pleszczyńska [29]. The remaining formulas were obtained by direct calculation, which was sometimes simple and sometimes complex. Note also that the models M_1 and M_6 are based on the definition of a bivariate Student t distribution as given in Demarta and McNeil [12]. In particular, in the case of M_1 we applied their formula (1) with $\mu = (0,0)$ and Σ with leading diagonal $(1,1)$ and off diagonal elements equal to ρ. In the case of M_6, we used their formula (2) with $\mu = (0,0), \gamma = (\gamma_1, \gamma_2), g(w) = w$ and $Z = (Z_1, Z_2)$, where Z_1 and Z_2 are independent standard normal.

Model $M_1 : \rho \in (-1,1)$,

$$\mu_{Y,X}(p) = \rho.$$

Model $M_2 : \theta \in \mathbb{R}$,

$\mu_{Y,X}(p) = \frac{\theta}{\sqrt{\theta^2+1}}$.

Model $M_3 : \rho \in (-1, 0]$,

$$\mu_{Y,X}(p) = \frac{\int_{1-p}^{p} 2\Phi\left(\frac{\rho\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right)+e^{\frac{1+2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{\rho^2-\rho\Phi^{-1}(t)-1}{\sqrt{1-\rho^2}}\right)-e^{\frac{1-2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{\rho^2+\rho\Phi^{-1}(t)-1}{\sqrt{1-\rho^2}}\right)\mathrm{d}t-p}{p+\sqrt{e}(\Phi(1-\Phi^{-1}(p))-1)}$$

for $p \le 0.5$, and

$$\mu_{Y,X}(p) = \frac{\int_{p}^{1} 2\Phi\left(\frac{\rho\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right)+e^{\frac{1+2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{\rho^2-\rho\Phi^{-1}(t)-1}{\sqrt{1-\rho^2}}\right)-e^{\frac{1-2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{\rho^2+\rho\Phi^{-1}(t)-1}{\sqrt{1-\rho^2}}\right)\mathrm{d}t-1+p}{1-p+\sqrt{e}(\Phi(\Phi^{-1}(p)+1)-1)}$$

for $p > 0.5$.

Model $M_4 : \rho \in (-1, 0]$,

$$\mu_{Y,X}(p) = \frac{\int_{1-p}^{p} 2\Phi\left(\frac{-\rho\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right)+e^{\frac{1+2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{-\rho^2+\rho\Phi^{-1}(t)+1}{\sqrt{1-\rho^2}}\right)-e^{\frac{1-2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{-\rho^2-\rho\Phi^{-1}(t)+1}{\sqrt{1-\rho^2}}\right)\mathrm{d}t-p}{\sqrt{e}(\Phi(1+\Phi^{-1}(p)))-p}$$

for $p \le 0.5$, and

$$\mu_{Y,X}(p) = \frac{\int_{p}^{1} 2\Phi\left(\frac{-\rho\Phi^{-1}(t)}{\sqrt{1-\rho^2}}\right)+e^{\frac{1+2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{-\rho^2+\rho\Phi^{-1}(t)+1}{\sqrt{1-\rho^2}}\right)-e^{\frac{1-2\rho\Phi^{-1}(t)-\rho^2}{2}}\Phi\left(\frac{-\rho^2-\rho\Phi^{-1}(t)+1}{\sqrt{1-\rho^2}}\right)\mathrm{d}t-1+p}{\sqrt{e}(\Phi(1-\Phi^{-1}(p)))-1+p}$$

for $p > 0.5$.

Model $M_5 : \rho \in (-1, 0]$,

$\mu_{Y,X}(p) = \frac{p-\Phi(\Phi^{-1}(p)-\rho)}{\Phi(\Phi^{-1}(p)+1)-p}$.

Model $M_6 : \gamma_1 < 0,\ \gamma_2 > 0$,

$\mu_{Y,X}(p) = \frac{l_{Y,X}(p)}{m_Y^+(p)}$, if $l_{Y,X}(p) \ge 0$, $\mu_{Y,X}(p) = \frac{l_{Y,X}(p)}{m_Y^-(p)}$, if $l_{Y,X}(p) < 0$, where $l_{Y,X}(p) = \frac{2p\gamma_1\sqrt{4+x_p^2}-\gamma_1 e^{\gamma_2\left(x_p-\sqrt{4+x_p^2}\right)}\left(x_p+\sqrt{4+x_p^2}\right)}{(1-p)\sqrt{4+x_p^2}}$, $m_Y^+(p) = \frac{e^{\gamma_1\left(y_p+\sqrt{4+y_p^2}\right)}\left(4-\gamma_1 y_p\left(4+y_p^2-y_p\sqrt{4+y_p^2}\right)\right)}{(1-p)\sqrt{4+y_p^2}(4+y_p^2)} - 2\gamma_1$, $m_Y^-(p) = -2\gamma_1\frac{p}{1-p} - \frac{e^{\gamma_1\left(y_{1-p}+\sqrt{4+y_{1-p}^2}\right)}\left(-4+\gamma_1 y_{1-p}\left(4+y_{1-p}^2-y_{1-p}\sqrt{4+y_{1-p}^2}\right)\right)}{(1-p)\sqrt{4+y_{1-p}^2}(4+y_{1-p}^2)}$, while $x_p = F^{-1}(p)$, $y_p = G^{-1}(p)$, $F(t) = \frac{e^{\gamma_2(t-\sqrt{4+t^2})}\left(6t+t^3+4\sqrt{4+t^2}+t^2\sqrt{4+t^2}+\gamma_2(8+2t^2+2t\sqrt{4+t^2})\right)}{2\sqrt{4+t^2}(4+t^2)}$, $G(t) = 1 + \frac{e^{\gamma_1(t+\sqrt{4+t^2})}\left(6t+t^3-4\sqrt{4+t^2}-t^2\sqrt{4+t^2}+\gamma_1(8+2t^2-2t\sqrt{4+t^2})\right)}{2\sqrt{4+t^2}(4+t^2)}$.

Model $M_7 : \theta \in [0, 1]$,

$\mu_{Y,X}(p) = \frac{(p-1)\theta\log(1-p)}{(1-\theta\log(1-p))p\log p}$.

Model $M_8 : \rho \in (-1, 0]$,

$\mu_{Y,X}(p) = \frac{p-\Phi(\Phi^{-1}(p)-2\rho)}{\Phi(\Phi^{-1}(p)+2)-p}$.

Model $M_9 : \theta \in [0, 3/4]$,

$\mu_{Y,X}(p) = \frac{2p(2p-1)(1-p)}{1-p+\frac{3}{10}\theta-\frac{1}{2}-3h(\theta,p)}$ for $p \in \left(0, \frac{\theta}{3}\right]$, $\mu_{Y,X}(p) = \frac{4\theta p(2p-1)(1-p)}{\frac{4}{15}\theta^2+3p(1-p)}$ for $p \in \left[\frac{\theta}{3}, 1-\frac{\theta}{3}\right]$, $\mu_{Y,X}(p) = \frac{2p(2p-1)(p-1)}{1-p+\frac{3}{10}\theta-\frac{1}{2}-3h(\theta,p)}$ for $p \in \left[1-\frac{\theta}{3}, 1\right)$, where $g(x) = \frac{2}{3}x^3 - x^2$ for $x \in [0,1]$, while $h(\theta,p) = \frac{1}{2}\left(g^{-1}\left(\frac{p}{\theta}-\frac{1}{3}\right)\right)^2 - \frac{1}{3}\left(g^{-1}\left(\frac{p}{\theta}-\frac{1}{3}\right)\right)^3 + \frac{\theta}{2}\left(g^{-1}\left(\frac{p}{\theta}-\frac{1}{3}\right)\right)^4 - \frac{2}{5}\theta\left(g^{-1}\left(\frac{p}{\theta}-\frac{1}{3}\right)\right)^5$.

Model $M_{10}: \eta \in [0, 1/2]$, $\delta > 0$, $\rho \in (-1, 1)$,

$\mu_{Y,X}(p) = -\rho + \frac{\eta\delta(1+\rho)[\Phi(z_p+\delta)-\Phi(z_p-\delta)]}{(1-2\eta)\varphi(z_p)+\eta[\varphi(z_p-\delta)+\varphi(z_p+\delta)+\delta\Phi(z_p+\delta)-\delta\Phi(z_p-\delta)]}$,
where $I(z) = (1-2\eta)\Phi(z) + \eta\Phi(z-\delta) + \eta\Phi(z+\delta)$, $z_p = I^{-1}(p)$, $\varphi(x) = \Phi'(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$.

Model $M_{11}: \theta \in [-1, 0]$,

$\mu_{Y,X}(p) = \frac{\sqrt{-8p\theta}+3\theta}{-3\theta(1-p)}$ for $p \in \left(0, -\frac{\theta}{2}\right]$, $\mu_{Y,X}(p) = \frac{\theta}{6p(1-p)}$ for $p \in \left(-\frac{\theta}{2}, 1+\frac{\theta}{2}\right]$, $\mu_{Y,X}(p) = \frac{\sqrt{-8(1-p)\theta}+3\theta}{-3\theta p}$ for $p \in \left(1+\frac{\theta}{2}, 1\right)$.

Model $M_{12}: \theta \geq 0$,

$\mu_{Y,X}(p) = \frac{p-B_{(2\theta+1,2\theta+1)}\left(B^{-1}_{(\theta+1,\theta+1)}(p)\right)}{B_{(3/2,2\theta)}\left(B^{-1}_{(3/2,\theta)}(1-p)\right)-1+p}$ for $p \in \left(0, \frac{1}{2}\right)$, $\mu_{Y,X}(p) = \frac{B_{(2\theta+1,2\theta+1)}\left(B^{-1}_{(\theta+1,\theta+1)}(1-p)\right)-1+p}{B_{(3/2,2\theta)}\left(B^{-1}_{(3/2,\theta)}(p)\right)-p}$ for $p \in \left[\frac{1}{2}, 1\right)$.

Appendix C. Auxiliary simulation results

Table C.1. Empirical significance levels and powers for $n = 200$; $\alpha = 0.05$.

M_1

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.044	0.033	0.029	0.048	0
$R^-(\epsilon)$	0.052	0.044	0.034		
$S^-(\epsilon)$	0.493	0.425	0.321	0.708	-0.20
$R^-(\epsilon)$	0.637	0.602	0.560		

M_2

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.040	0.041	0.057	0.049	0
$R^-(\epsilon)$	0.044	0.043	0.045		
$S^-(\epsilon)$	0.749	0.690	0.603	0.791	-0.20
$R^-(\epsilon)$	0.793	0.781	0.776		

M_3

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.053	0.049	0.053	0.055	0
$R^-(\epsilon)$	0.048	0.049	0.053		
$S^-(\epsilon)$	0.701	0.621	0.536	0.796	-0.20
$R^-(\epsilon)$	0.793	0.784	0.776		

M_4

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.042	0.043	0.041	0.046	0
$R^-(\epsilon)$	0.046	0.043	0.042		
$S^-(\epsilon)$	0.676	0.653	0.597	0.725	-0.20
$R^-(\epsilon)$	0.723	0.711	0.692		

M_5

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.076	0.073	0.079	0.038	0
$R^-(\epsilon)$	0.041	0.042	0.040		
$S^-(\epsilon)$	0.614	0.584	0.563	0.623	-0.20
$R^-(\epsilon)$	0.635	0.621	0.604		

M_6

ϵ	0.100	0.075	0.050	T	c
$S^-(\epsilon)$	0.041	0.040	0.032	0.049	0
$R^-(\epsilon)$	0.042	0.044	0.042		
$S^-(\epsilon)$	0.529	0.475	0.374	0.707	0.18
$R^-(\epsilon)$	0.687	0.669	0.638		

M_7

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.062	0.072	0.076	0.046	0.0
$R^-(\epsilon)$	0.049	0.049	0.050		
$S^-(\epsilon)$	0.668	0.653	0.630	0.593	0.20
$R^-(\epsilon)$	0.671	0.666	0.660		

M_8

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.076	0.085	0.091	0.020	0
$R^-(\epsilon)$	0.034	0.034	0.032		
$S^-(\epsilon)$	0.722	0.704	0.680	0.574	-0.35
$R^-(\epsilon)$	0.649	0.631	0.615		

M_9

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.045	0.051	0.053	0.054	0
$R^-(\epsilon)$	0.055	0.058	0.055		
$S^-(\epsilon)$	0.845	0.793	0.700	0.689	0.40
$R^-(\epsilon)$	0.830	0.825	0.805		

M_{10}

ϵ	0.100	0.075	0.050	T	ρ/δ
$S^-(\epsilon)$	0.055	0.041	0.040	0.060	0/0
$R^-(\epsilon)$	0.058	0.058	0.058		
$S^-(\epsilon)$	0.811	0.867	0.918	0.644	0.20/2.5
$R^-(\epsilon)$	0.752	0.824	0.886		

M_{11}

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.058	0.051	0.040	0.054	0
$R^-(\epsilon)$	0.058	0.057	0.053		
$S^-(\epsilon)$	0.746	0.857	0.941	0.490	-0.15
$R^-(\epsilon)$	0.662	0.746	0.843		

M_{12}

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.060	0.045	0.049	0.059	0
$R^-(\epsilon)$	0.057	0.056	0.052		
$S^-(\epsilon)$	0.906	0.936	0.951	0.452	0.50
$R^-(\epsilon)$	0.780	0.817	0.841		

TABLE C.2. Empirical significance levels and powers for $n = 400$; $\alpha = 0.05$.

M_1

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.051	0.036	0.030	0.050	0
$R^-(\epsilon)$	0.059	0.052	0.046		
$S^-(\epsilon)$	0.565	0.478	0.375	0.705	-0.150
$R^-(\epsilon)$	0.646	0.631	0.596		

M_2

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.050	0.059	0.060	0.042	0
$R^-(\epsilon)$	0.040	0.042	0.041		
$S^-(\epsilon)$	0.717	0.692	0.619	0.792	-0.140
$R^-(\epsilon)$	0.789	0.791	0.784		

M_3

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.048	0.053	0.048	0.056	0
$R^-(\epsilon)$	0.052	0.058	0.054		
$S^-(\epsilon)$	0.642	0.581	0.481	0.755	-0.135
$R^-(\epsilon)$	0.747	0.740	0.728		

M_4

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.048	0.044	0.041	0.058	0
$R^-(\epsilon)$	0.052	0.052	0.049		
$S^-(\epsilon)$	0.691	0.668	0.626	0.726	-0.144
$R^-(\epsilon)$	0.721	0.723	0.702		

M_5

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.063	0.069	0.072	0.042	0
$R^-(\epsilon)$	0.043	0.043	0.042		
$S^-(\epsilon)$	0.571	0.540	0.503	0.637	-0.144
$R^-(\epsilon)$	0.641	0.626	0.611		

M_6

ϵ	0.100	0.075	0.050	T	c
$S^-(\epsilon)$	0.049	0.046	0.040	0.058	
$R^-(\epsilon)$	0.052	0.048	0.051		0
$S^-(\epsilon)$	0.538	0.496	0.414	0.668	0.108
$R^-(\epsilon)$	0.662	0.659	0.641		

M_7

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.066	0.071	0.066	0.048	0
$R^-(\epsilon)$	0.050	0.051	0.054		
$S^-(\epsilon)$	0.718	0.698	0.662	0.667	0.144
$R^-(\epsilon)$	0.739	0.747	0.743		

M_8

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.078	0.083	0.087	0.021	0
$R^-(\epsilon)$	0.036	0.035	0.035		
$S^-(\epsilon)$	0.698	0.692	0.673	0.595	-0.250
$R^-(\epsilon)$	0.648	0.638	0.625		

M_9

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.054	0.060	0.062	0.059	0
$R^-(\epsilon)$	0.054	0.054	0.060		
$S^-(\epsilon)$	0.870	0.853	0.790	0.722	0.280
$R^-(\epsilon)$	0.845	0.852	0.847		

M_{10}

ϵ	0.100	0.075	0.050	T	ρ/δ
$S^-(\epsilon)$	0.059	0.052	0.044	0.055	0/0
$R^-(\epsilon)$	0.057	0.059	0.059		
$S^-(\epsilon)$	0.792	0.842	0.874	0.597	0.15/1.94
$R^-(\epsilon)$	0.724	0.776	0.822		

M_{11}

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.055	0.058	0.056	0.048	0
$R^-(\epsilon)$	0.052	0.054	0.056		
$S^-(\epsilon)$	0.731	0.850	0.944	0.490	-0.106
$R^-(\epsilon)$	0.643	0.743	0.850		

M_{12}

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.060	0.052	0.053	0.060	0
$R^-(\epsilon)$	0.054	0.054	0.055		
$S^-(\epsilon)$	0.942	0.961	0.973	0.533	0.300
$R^-(\epsilon)$	0.839	0.876	0.912		

TABLE C.3. Empirical significance levels and powers for $n = 1000$; $\alpha = 0.05$.

M_1

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.053	0.061	0.053	0.055	0
$R^-(\epsilon)$	0.052	0.050	0.047		
$S^-(\epsilon)$	0.558	0.479	0.390	0.683	-0.093
$R^-(\epsilon)$	0.654	0.625	0.600		

M_2

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.043	0.048	0.052	0.047	0
$R^-(\epsilon)$	0.048	0.046	0.042		
$S^-(\epsilon)$	0.749	0.717	0.627	0.806	-0.090
$R^-(\epsilon)$	0.794	0.795	0.787		

M_3

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.041	0.045	0.059	0.045	0
$R^-(\epsilon)$	0.042	0.042	0.043		
$S^-(\epsilon)$	0.706	0.633	0.531	0.775	-0.088
$R^-(\epsilon)$	0.776	0.777	0.770		

M_4

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.056	0.055	0.051	0.060	0
$R^-(\epsilon)$	0.058	0.057	0.058		
$S^-(\epsilon)$	0.752	0.742	0.724	0.775	-0.093
$R^-(\epsilon)$	0.777	0.773	0.774		

M_5

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.064	0.054	0.063	0.042	0
$R^-(\epsilon)$	0.050	0.054	0.057		
$S^-(\epsilon)$	0.528	0.506	0.479	0.616	-0.091
$R^-(\epsilon)$	0.624	0.596	0.595		

M_6

ϵ	0.100	0.075	0.050	T	c
$S^-(\epsilon)$	0.050	0.039	0.048	0.050	
$R^-(\epsilon)$	0.049	0.047	0.049		0
$S^-(\epsilon)$	0.628	0.601	0.537	0.681	0.065
$R^-(\epsilon)$	0.678	0.685	0.696		

M_7

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.060	0.058	0.057	0.049	0
$R^-(\epsilon)$	0.056	0.052	0.052		
$S^-(\epsilon)$	0.713	0.685	0.677	0.678	0.088
$R^-(\epsilon)$	0.747	0.754	0.757		

M_8

ϵ	0.100	0.075	0.050	T	ρ
$S^-(\epsilon)$	0.067	0.075	0.082	0.037	0
$R^-(\epsilon)$	0.042	0.043	0.040		
$S^-(\epsilon)$	0.653	0.651	0.634	0.591	-0.170
$R^-(\epsilon)$	0.638	0.632	0.619		

M_9

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.051	0.051	0.058	0.048	0
$R^-(\epsilon)$	0.048	0.050	0.056		
$S^-(\epsilon)$	0.872	0.859	0.784	0.730	0.170
$R^-(\epsilon)$	0.862	0.865	0.850		

M_{10}

ϵ	0.100	0.075	0.050	T	ρ/δ
$S^-(\epsilon)$	0.057	0.043	0.046	0.044	0/0
$R^-(\epsilon)$	0.044	0.041	0.041		
$S^-(\epsilon)$	0.742	0.771	0.778	0.576	0.10/1.47
$R^-(\epsilon)$	0.685	0.711	0.741		

M_{11}

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.050	0.053	0.057	0.046	0
$R^-(\epsilon)$	0.043	0.043	0.046		
$S^-(\epsilon)$	0.742	0.848	0.952	0.487	-0.067
$R^-(\epsilon)$	0.667	0.737	0.865		

M_{12}

ϵ	0.100	0.075	0.050	T	θ
$S^-(\epsilon)$	0.059	0.052	0.048	0.058	0
$R^-(\epsilon)$	0.058	0.047	0.056		
$S^-(\epsilon)$	0.866	0.904	0.937	0.476	0.145
$R^-(\epsilon)$	0.761	0.806	0.845		

In the simulation study reported above $r = 2000$ Monte Carlo replications were used to approximate p-values, and each experiment is repeated 1000 times to compare empirical significance levels and powers at the nominal significance level $\alpha = 0.05$.

Acknowledgements. The authors would like to thank an Associate Editor for several useful suggestions and two reviewers for their helpful comments on the first version of this paper. The authors also wish to express their gratitude to Tadeusz Inglot for helpful conversations, as well as Jan Mielniczuk, Yakov Nikitin and Sonia Petrone for help in obtaining papers related to the topic. The research of the first author was co-founded by the European Union by the European Social Fund *via* the Project entitled "PO KL Information Technologies: Research and Their Interdisciplinary Applications" (grant no. UDA-POKL.04.01.01-00-051/10-00).

References

[1] W. Albers, Stop-loss premiums under dependence. *Insurance: Math. Econom.* **24** (1999) 173−185.
[2] N. Balakrishnan and Ch.-D.Lai, Continuous Bivariate Distributions. Springer, Dordrecht (2009).
[3] G.F. Barrett and S.G. Donald, Consistent tests for stochastic dominance. *Econometrica* **71** (2003) 71–104.
[4] D.G. Baur, The structure and degree of dependence: A quantile regression approach. *J. Bank. Fin.* **37** (2013) 786–798.
[5] T. Bednarski and T. Ledwina, A note on the weak convergence of an estimator of monotonic dependence function of two random variables. *Math. Oper. Stat., Ser. Statist.* **13** (1982) 555–560.
[6] G.D. Berentsen and D. Tjøstheim, Recognizing and visualizing departures from independence in bivariate data using local Gaussian correlation. *Statist. Comput.* **24** (2014) 785–801.
[7] P. Billingsley, Convergence of Probability Measures. Wiley, New York (1968).
[8] R. Cont, Empirical properties of asset returns: stylized facts and statistical issues. *Quant. Fin.* **1** (2001) 223–236.
[9] M. Csörgő, J.L. Gastwirth and R. Zitikis, Asymptotic confidence bands for the Lorenz and Bonferroni curves based on the empirical Lorenz curve. *J. Stat. Plan. Inference* **74** (1998) 65–91.
[10] J. Ćwik, T. Kowalczyk, A. Kowalski, E. Pleszczyńska, W. Szczęsny and T.Wierzbowska, Monotonic dependence function: background, new results and applications. In Mathematical Statistics and Probability Theory. Vol. 2 of *Lect. Notes Statistics.* Springer, New York (1980) 111–121.
[11] Yu. Davydov and V. Egorov, Functional limit theorems for induced order statistics. *Math. Methods Statist.* **9** (2000) 297–313.
[12] S. Demarta and A.J. McNeil, The t copula and related copulas. *Inter. Statist. Rev.* **73** (2005) 111–129.
[13] M. Denuit, R.J. Huang and L.Y. Tzeng, Almost expectation and excess dependence notions. *Theory Decision* **79** (2015) 375–401.
[14] M. Denuit and M. Mesfioui, Preserving the Rothschild−Stiglitz type increase in risk with background risk: A characterization. *Insur.: Math. Econ.* **72** (2017) 1–5.
[15] M. Denuit and O. Scaillet, Nonparametric tests for positive quadrant dependence. *J. Fin. Econ.* **2** (2004) 422–450.
[16] J. Dhaene and M.J. Goovaerts, On the dependency of risks in the individual life model. *Insur.: Math. Econ.* **19** (1997) 243–253.
[17] M. Egozcue, L.F. Garcia, W.-K. Wong and R. Zitikis, Grüss-type bounds for covariances and the notion of quadrant dependence in expectation. *Central Europ. J. Math.* **9** (2011) 1288–1297.
[18] P.A. Embrechts, A. McNeil and D. Straumann, Correlation and dependency in risk management: properties and pitfalls. In Risk Management: Value at Risk and Beyond, edited by M. Dempster and H. Moffatt. Cambridge University Press (2002) 176–223.
[19] Ch. M. Goldie, Convergence theorems for empirical Lorenz curves and their inverses. *Adv. Appl. Probab.* **9** (1977) 765–791.
[20] I. Gijbels, M. Omelka and D. Sznajder, Positive quadrant dependence tests for copulas. *Canadian J. Statist.* **38** (2010) 555–581.
[21] X. Guo and J. Li, Confidence bands for expectation dependence with applications. *Insur.: Math. Econ.* **68** (2016) 141–149.
[22] X. Guo, J.Y. Li, D.R. Liu and J.L. Wang, Preserving the Rothschild−Stiglitz type of increasing risk with background risk. *Insur.: Math. Econ.* **70** (2016) 144–149.
[23] M.S. Handcock and M. Morris, Relative Distribution Methods in the Social Sciences. Springer, New York (1999).
[24] S.K. Hong, K.O. Lew and R. MacMinn, Brockett, P. Mossin's theorem given random initial wealth. *J. Risk Insur.* **78** (2011) 309–324.
[25] R. Kaas, How to (and how not to) compute stop-loss premiums in practice. *Insur.: Math. Econ.* **13** (1993) 241–254.
[26] T. Kowalczyk, General definition and sample counterparts of monotonic dependence functions of bivariate distributions. *Math. Oper. Stat., Ser. Statist.* **8** (1977) 351–365.
[27] Kowalczyk, T. Shape of the monotonic dependence function. *Math. Oper. Stat., Ser. Statist.* **13** (1982) 183–192.
[28] T. Kowalczyk, A. Kowalski, A. Matuszewski and E. Pleszczyńska, Applications of monotonic dependence functions: Descriptive aspects. *Zastosow. Mat.* **XVI** (1980) 579–591.
[29] T. Kowalczyk and E. Pleszczyńska, Monotonic dependence functions of bivariate distributions. *Ann. Statist.* **5** (1977) 1221–1227.
[30] T. Kowalczyk and W. Szczesny, Preliminary concepts of bivariate dependence. In Grade Models and Methods for Data Analysis, edited by T. Kowalczyk, E. Pleszczyńska, F. Ruland (2004) 185–215.

[31] M. Ledoux and M. Talagrand, Un critère sur les petites boules dans le théorème limite central. *Probab. Theory Relat. Fields* **77** (1988) 29–47.

[32] M. Ledoux and M. Talagrand, Probability in Banach Spaces. Springer, Berlin (1991).

[33] T. Ledwina, Visualizing association structure in bivariate copulas using new dependence function. In Stochastic Models, Statistics and Their Applications. Vol. 122 of *Springer Proc. Math. Statist.*, edited by A. Steland, E. Rafajłowicz and K. Szajowski (2015) 19–27.

[34] T. Ledwina and G. Wyłupek, Validation of positive quadrant dependence. *Insur.: Math. Econ.* **56** (2014) 38–47.

[35] R. Li, Y. Cheng and J.P. Fine, Quantile association regression models. *J. Amer. Stat. Assoc.* **109** (2014) 230–242.

[36] O. Linton and Y.-J. Whang, The quantilogram: with an application to evaluating directional predictability. *J. Econ.* **141** (2007) 250–282.

[37] P. Muliere and S. Petrone, Generalized Lorenz curve and monotone dependence orderings. *Metron* L (1992) 19–38.

[38] K.R.M. Nair and N. Sreelakshmi, The new Zenga curve in the context of reliability analysis. *Commun. Stat. Theory Methods* **45** (2016) 6540–6552.

[39] C.R. Rao and L.C. Zhao, Convergence theorems for empirical cumulative quantile regression functions. *Math. Methods Statist.* **4** (1995a) 81–91.

[40] C.R. Rao and L.C. Zhao, Correction. Convergence theorems for empirical cumulative quantile regression functions. *Math. Methods Statist.* **4** (1995b) 359.

[41] O. Scaillet, A Kolmogorov−Smirnov type test for positive quadrant dependence. *Canadian J. Statist.* **33** (2005) 415–427.

[42] E. Schechtman, A. Shelef, S. Ytzhaki and R. Zitikis, Testing hypotheses about absolute concentration curves and marginal conditional stochastic dominance. *Econom. Theory* **24** (2008) 1044–1062.

[43] S.M. Tse, On the cumulative quantile regression process. *Math. Methods Statist.* **18** (2009) 270–279.

[44] V.S. Tsirel'son, The density of the distribution of the maximum of a Gaussian process. *Theory of Probab. Appl.* **20** (1975) 847–856.

[45] R. Wright, Expectation dependence of random variables, with an application in portfolio theory. *Theory Decis.* **22** (1987) 111–124.

[46] T. Yanagimoto, Families of positively dependent random variables. *Ann. Inst. Statist. Math.* **24** (1972) 559–573.

[47] T. Yanagimoto and M. Sibuya, Stochastically larger component of a random vector. *Ann. Institute Statist. Math.* **24** (1972) 259–269.

[48] M. Zenga, Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. *Statist. Appl.* (2007) 3–27.

[49] X. Zhu, X. Guo, L. Lin and L. Zhu, Testing for positive expectation dependence. *Ann. Inst. Statist. Math.* **68** (2016) 135–153.