

An analysis of a heuristic procedure to evaluate tail (in)dependence

Marta Ferreira

Center of Mathematics of Minho University and Department of Mathematics and Applications of Minho University
Campus de Gualtar 4710 - 057 Braga, Portugal
msferreira@math.uminho.pt

Sérgio Silva

Department of Mathematics and Applications of Minho University
Campus de Gualtar 4710 - 057 Braga, Portugal
sjoaosilva@yahoo.com

Abstract

Measuring tail dependence is an important issue in many applied sciences in order to quantify the risk of simultaneous extreme events. A usual measure is given by the tail dependence coefficient. The characteristics of events behave quite differently as these become more extreme, whereas we are in the class of asymptotic dependence or in the class of asymptotic independence. The literature has emphasized the asymptotic dependent class but wrongly infer tail dependence will result in the over-estimation of extreme value dependence and consequently of the risk. In this paper we analyze this issue through simulation based on a heuristic procedure.

keywords: tail dependence, asymptotic independence, extreme value theory

AMS 2000 Subject Classification Primary: 60G70; Secondary: 62G32

1 Introduction

The degree of association between concurrent rainfall extremes at different locations may lead to a better understanding of extreme rainfall events, a very important matter due to their severe impacts on the economy and the environment. The globalization and an absence of market regulation increased the dependence between financial asset returns and thus the risk of simultaneous crashes. Pearson's correlation is not an appropriate measure of dependence whenever extreme realizations are important. It gives the same weight to extreme values as for all of the other observations and the dependence characteristics for extreme realizations may differ from all others in the sample. For more details, see e.g. Embrechts *et al.* ([5], 2002). The most used measure of

tail dependence is the so called tail dependence coefficient (TDC), a concept introduced by Sibuya ([18], 1960), which is defined as follows:

$$\lambda = \lim_{t \downarrow 0} P(F_2(X_2) > 1 - t | F_1(X_1) > 1 - t), \quad (1)$$

where F_1 and F_2 are the distribution functions (d.f.'s) of the random variables (r.v.'s) X_1 and X_2 , respectively, which are considered continuous. Observe that the TDC can also be formulated through the copula function introduced by Sklar ([19], 1959). A copula function C is a d.f. whose marginals are standard uniform, i.e., if C is the copula function of (X_1, X_2) , having joint d.f. F , then

$$F(x_1, x_2) = P(F_1(X_1) \leq F_1(x_1), F_2(X_2) \leq F_2(x_2)) = C(F_1(x_1), F_2(x_2))$$

and thus

$$\lambda = 2 - \lim_{t \downarrow 0} \frac{1 - C(1 - t, 1 - t)}{t}. \quad (2)$$

If $0 < \lambda \leq 1$, the r.v.'s X_1 e X_2 are said to be tail dependent whose degree of dependence is measured through λ ($\lambda = 1$ means total dependence in the tail). The case $\lambda = 0$, corresponds to asymptotic independence in the tail. However, as noticed in Ledford and Tawn ([12, 13], 1996/1997), it may occur a residual tail dependence captured through the convergence rate of $P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t)$ towards zero, as $t \downarrow 0$. More precisely, considering

$$P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t) = t^{1/\eta} L(t), \quad (3)$$

where L is a slow varying function at 0, i.e., $L(tx)/L(t) \rightarrow 1$, as $t \downarrow 0$ and $x > 0$. The parameter $\eta \in (0, 1]$, known as Ledford and Tawn coefficient, measures the residual tail dependence and the function L the relative strength of dependence given a particular value of η . Observe that $\eta = 1$ and $L(t)$ converging to some positive constant corresponds to tail dependence ($\lambda > 0$), whilst $\eta < 1$ means tail independence. If $\eta = 1/2$ we have (almost) perfect independence (perfect if $L(t) = 1$), and for $0 < \eta < 1/2$ or $1/2 < \eta < 1$ we have association, respectively, negative (i.e., $P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t) < P(F_1(X_1) > 1 - t)P(F_2(X_2) > 1 - t)$) or positive (i.e., $P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t) > P(F_1(X_1) > 1 - t)P(F_2(X_2) > 1 - t)$).

Relation (3) also means that the function $q(t) = P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t)$ is regularly varying (of first order) with index $1/\eta$. In Draisma *et al.* ([4], 2004), it is considered a refinement of this relation under a second order regularly varying condition. More precisely, it is assumed that the limit

$$\lim_{t \downarrow 0} \left(\frac{P(F_1(X_1) > 1 - tx, F_2(X_2) > 1 - ty)}{q(t)} - c(x, y) \right) / q_1(t) = c_1(x, y) \quad (4)$$

exists, for all $x, y \geq 0$ and $x + y > 0$, with $q_1(t) \rightarrow 0$, as $t \downarrow 0$, being a regularly varying function of index $\tau \geq 0$ and c_1 a non-constant function and non-multiple of c . It is also assumed that the convergence is uniform in $\{(x, y) \in [0, \infty[^2: x^2 + y^2 = 1\}$, that $l = \lim_{t \downarrow 0} q(t)/t$ exists and, without loss of generality, $c(1, 1) = 1$. In addition, the function c is homogeneous of order $1/\eta$, i.e., $c(tx, ty) = t^{1/\eta} c(x, y)$.

Now observe that,

$$P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t) = P\left(W_1 > \frac{1}{t}, W_2 > \frac{1}{t}\right),$$

with $W_j = (1 - F_j(X_j))^{-1}$, $j = 1, 2$, and hence we can write

$$P(W_1 > t, W_2 > t) = t^{-1/\eta} L(t^{-1}). \quad (5)$$

Therefore, η corresponds to the tail index of

$$T = \min(W_1, W_2), \quad (6)$$

and thus it can be estimated as so. This will be addressed in Section 2.

An alternative measure for the residual tail dependence was introduced in Coles *et al.* ([3], 1999). By considering

$$P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t) = [P(F_1(X_1) > 1 - t)P(F_2(X_2) > 1 - t)]^{1/(2\eta)},$$

and applying logarithms to both members, we derive

$$\frac{1}{2\eta} = \frac{\log P(F_1(X_1) > 1 - t, F_2(X_2) > 1 - t)}{\log P(F_1(X_1) > 1 - t) + \log P(F_2(X_2) > 1 - t)},$$

or

$$\bar{\chi} = 2\eta - 1 = \frac{2 \log t}{\log \bar{C}(1 - t, 1 - t)} - 1, \quad (7)$$

with $\bar{\chi} \in [-1, 1]$ and \bar{C} corresponding to the survival copula, i.e.,

$$\bar{F}(x_1, x_2) = P(F_1(X_1) > F_1(x_1), F_2(X_2) > F_2(x_2)) = \bar{C}(F_1(x_1), F_2(x_2)).$$

Observe that $\bar{\chi} < 1$ means tail independence ($\lambda = 0$) and if $\bar{\chi} = 1$, we have tail dependence ($\lambda > 0$). We also have positive and negative association whenever $\bar{\chi} > 0$ and $\bar{\chi} < 0$, respectively, with $\bar{\chi} = 0$ corresponding to (almost) exact independence. Estimators for λ and $\bar{\chi}$ based on the expressions (2) and (7), respectively, will be also presented in Section 2.

The behavior of events within the class of asymptotic dependence is quite different from the one detected in the class of asymptotic independence. Both forms allow dependence between moderately large values of each variable, but only when the variables exhibit tail dependence the very largest values from each variable can occur together. If we wrongly infer tail dependence, an over-estimation of the extreme value dependence takes place and consequently of the risk. This over-estimation is related to the degree of residual dependence which is measured through η or $\bar{\chi}$. Therefore, it is important to assess whereas a data set presents tail dependence or independence and to quantify the degree of dependence for the appropriate dependence class. This can be done through the estimation of λ and of η (or $\bar{\chi}$) together with tests for tail independence. These topics can be found in many references such as Huang ([9], 1992), Joe ([10], 1997), Coles *et al.* ([3], 1999), Frahm *et al.* ([7], 2005) and Schmidt and

Stadtmüller ([17], 2006) for the TDC estimation, Ledford and Tawn ([12, 13], 1996/1997) and Peng ([15], 1999) concerning the estimation of η and Coles *et al.* ([3], 1999) for the $\bar{\chi}$ estimation. The tail independence tests can be seen in, e.g., Poon *et al.* ([16], 2004) and Draisma *et al.* ([4], 2004).

Most of the non parametric estimation of extremal parameters requires the choice of the number k of upper order statistics to be used in it. A paradigmatic example is the univariate tail index estimation of regularly varying distributions (for a survey, see Beirlant *et al.* [1] 2012 and references therein). A similar problem exists for tail dependence estimation. In practice, we have to deal with a trade-off between variance and bias, since small values of k correspond to larger variance whilst large values of k increase the bias of the estimators. Figure 1 illustrates this issue. Observe that the true value (horizontal line) can be inferred from a kind of first stability region within the sample path of estimators. In order to overcome this problem, Frahm *et al.* ([7], 2005) developed a heuristic procedure where k is estimated based on a simple plateau-finding algorithm after smoothing the latter plot by some box kernel. They proposed some values for the bandwidth but no study was carried out in order to evaluate possible choices. In this paper we address this issue through a simulation study, by applying the heuristic procedure to nonparametric estimators of the TDC. In addition, we also analyze the performance of the procedure when applied to the estimation of η and $\bar{\chi}$, as well as, within the context of the referred tail independence tests (Section 3). An illustration with financial data is presented in Section 4. We end with some final remarks (Section 5).

2 Inference on the extremal (in)dependence

Consider $(X_1^{(1)}, X_2^{(1)}), \dots, (X_1^{(n)}, X_2^{(n)})$ independent and identically distributed (i.i.d.) copies of the random pair (X_1, X_2) . From (2), it is possible to deduce the estimator (Joe *et al.*, [11] 1992):

$$\widehat{\lambda}^{\text{SEC}} = 2 - \frac{1 - \widehat{C}(1 - k/n, 1 - k/n)}{k/n}, \quad 1 \leq k < n. \quad (8)$$

By using $\log(1 - t) \sim -t$, with $t \approx 0$, it can be derived the estimator (Coles *et al.* 1999):

$$\widehat{\lambda}^{\text{LOG}} = 2 - \frac{\log \widehat{C}(1 - k/n, 1 - k/n)}{\log(1 - k/n)}, \quad 1 \leq k < n, \quad (9)$$

where \widehat{C} denotes the empirical copula given by

$$\widehat{C}(1 - k/n, 1 - k/n) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{F_1(X_1^{(i)}) \leq 1 - \frac{k}{n}, F_2(X_2^{(i)}) \leq 1 - \frac{k}{n}\}}, \quad 1 \leq k < n, \quad (10)$$

with $\mathbb{1}$ denoting the indicator function. If the marginal d.f.'s F_j , $j = 1, 2$, are unknown, we can replace them by the empirical d.f.'s \widehat{F}_j , $j = 1, 2$, respectively. For more accurate estimates, it is considered

$$\widehat{F}_j(u) = \frac{1}{n+1} \sum_{i=1}^n \mathbb{1}_{\{X_j^{(i)} \leq u\}}, \quad j = 1, 2. \quad (11)$$

See Beirlant *et al.* ([2], 2004; Section 9.4.1) for more details. Note that both estimators depend on the parameter k , the number of upper order statistics involved in the estimation. The choice of k is of major difficulty within these estimators because of the compromise between variance and bias explained in the introduction. To ensure properties as asymptotic normality and consistency it is necessary to assume that $k \equiv k_n$ is an intermediate sequence, i.e.,

$$k \rightarrow \infty \quad \text{and} \quad k/n \rightarrow 0, \quad \text{as } n \rightarrow \infty$$

(see Huang [9] 1992 and Schmidt and Stadtmüller [17] 2006).

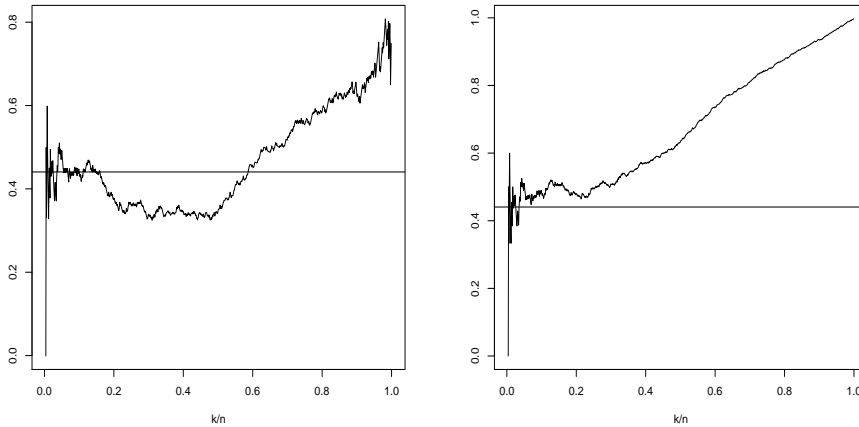


Figure 1: Sample path of estimator $\widehat{\lambda}^{\text{LOG}}$ (left) and estimator $\widehat{\lambda}^{\text{SEC}}$ (right), plotted against k/n , $1 \leq k < n$, considering $n = 1000$ realizations of a bivariate Student- t . The horizontal lines correspond to the true values.

We have already seen that, by considering (5), coefficient η corresponds to the tail index of the r.v. T defined in (6). The tail index estimation has been largely exploited in literature and a survey on this topic can be seen in, e.g., Beirlant *et al.* ([2], 2004). The most used estimator within positive tail indexes is the Hill estimator (Hill [8] 1975). More precisely, considering in (6) the respective empirical counterparts, we have

$$T_i^{(n)} = \min \left(\widehat{W}_{1,i}, \widehat{W}_{2,i} \right), \quad i = 1, \dots, n \quad (12)$$

with $\widehat{W}_{j,i} = (1 - \widehat{F}_j(X_{j,i}))^{-1}$ and \widehat{F}_j given in (11), $j = 1, 2$, $i = 1, \dots, n$. Thus, considering the order statistics, $T_{n:n}^{(n)} \geq T_{n:n-1}^{(n)} \geq \dots \geq T_{n:n-k}^{(n)}$, the Hill estimator for coefficient η is given by

$$\widehat{\eta} = \frac{1}{k} \sum_{i=1}^k \log \frac{T_{n:n-i+1}^{(n)}}{T_{n:n-k}^{(n)}}, \quad 1 \leq k < n. \quad (13)$$

Observe that $\widehat{\eta}$ is also a function of the parameter k , under the same conditions described above and thus, suffering from the same problem involving the bias and variance.

Observe that, from the first equality in (7), we can derive the estimator

$$\widetilde{\chi} = 2\widehat{\eta} - 1,$$

with $\widehat{\eta}$ given in (13). From the second equality in (7), it is obtained the estimator (Coles *et al.* 1999)

$$\widehat{\chi} = \frac{2 \log(k/n)}{\log \widehat{C}(1 - k/n, 1 - k/n)} - 1, \quad 1 \leq k < n, \quad (14)$$

where \widehat{C} denotes the empirical survival copula,

$$\widehat{C}(1 - k/n, 1 - k/n) = \frac{1}{n} \sum_{i=1}^n \mathbb{1}_{\{F_1(X_1^{(i)}) > 1 - \frac{k}{n}, F_2(X_2^{(i)}) > 1 - \frac{k}{n}\}}, \quad 1 \leq k < n. \quad (15)$$

In case the margins are unknown, we can replace F_j by \widehat{F}_j , $j = 1, 2$, given in (11). Once again, we have dependency on the parameter k .

In most of the cases, the TDC estimators do not behave well under asymptotic independence, i.e., whenever $\lambda = 0$ (see, e.g., Frahm *et al.* [7] 2005 and Ferreira [6] 2013). A possible way to deal with this problem is to consider preliminar tests for tail independence. Poon *et al.* [16] (2002) suggest to test $H_0 : \eta = 1$ versus $H_1 : \eta < 1$, that is, dependence versus independence, based on estimator $\widehat{\eta}$ in (13). Considering $k \equiv k_n$ an intermediate sequence and under some quite general additional conditions, we have $\sqrt{k}(\eta - \widehat{\eta})$ approximately $N(0, \eta^2)$, as $n \rightarrow \infty$. Thus, we reject H_0 in favor of H_1 , at the significance level α , if

$$\widehat{\eta} + z_{1-\alpha} \frac{\widehat{\eta}}{\sqrt{k}} < 1 \quad (16)$$

where $z_{1-\alpha}$ denotes the $(1 - \alpha)$ -quantile of $N(0, 1)$.

An analogous test was developed in Draisma *et al.* ([4], 2004), based on relation (4). More precisely, assuming that (4) holds for a function c with first derivatives $c_x = \partial c(x, y)/\partial x$ and $c_y = \partial c(x, y)/\partial y$, and considering $k \equiv k_n$ an intermediate sequence such that $\sqrt{k}q_1(q^{-1}(k/n)) \rightarrow 0$, with $n \rightarrow \infty$, then $\sqrt{k}(\widehat{\eta} - \eta)$ is asymptotically Normal with null mean value and variance

$$\sigma^2 = \eta^2(1 - l)(1 - 2lc_x(1, 1)c_y(1, 1)).$$

Consider

$$\widehat{l} = \frac{k}{n} T_{n:n-k}^{(n)}, \quad \widehat{c}_x(1, 1) = \frac{\widehat{k}^{5/4}}{n} \left(T_{n:n-k}^{(n, \widehat{k}^{5/4})} - T_{n:n-k}^{(n)} \right)$$

with $\widehat{k} = k/\widehat{l}$ and $T_{n:i}^{(n,u)}$, $i = 1, \dots, n$, the ordinal statistics of

$$T_i^{(n,u)} = \min \left(\widehat{W}_{1,i}(1+u), \widehat{W}_{2,i} \right), \quad i = 1, \dots, n.$$

Defining similarly $\widehat{c}_y(1, 1)$, if (4) holds under the above mentioned conditions, then $\widehat{l} \xrightarrow{P} l$, where \xrightarrow{P} denotes convergence in probability. Moreover, if $\eta = 1$, then

$$\widehat{c}_x(1, 1) \xrightarrow{P} c_x(1, 1), \quad \widehat{c}_y(1, 1) \xrightarrow{P} c_y(1, 1) \quad \text{and} \quad \widehat{\sigma} \xrightarrow{P} \sigma,$$

where

$$\hat{\sigma}^2 = \hat{\eta}^2(1 - \hat{l})(1 - 2\hat{l} \hat{c}_x(1, 1)\hat{c}_y(1, 1)). \quad (17)$$

with $\hat{\eta}$ corresponding to the Hill estimator of η , given in (13). Therefore, for the same test hypotheses, we reject H_0 if

$$\hat{\eta} + z_{1-\alpha} \frac{\hat{\eta}}{\sqrt{k}} \sqrt{(1 - \hat{l})(1 - 2\hat{l} \hat{c}_x(1, 1)\hat{c}_y(1, 1))} < 1. \quad (18)$$

Observe that the variance in test (18) includes a correctness factor when compared with the one in (16). This will render its value slightly smaller, making the test more accurate under tail independence, as shall be seen in the simulations afterwards.

3 Simulations

In this section we analyze the ‘‘plateau-finding’’ heuristic procedure presented in Frahm *et al.* ([7], 2005). A stability on the sample path of the graph $(k, \hat{\lambda}(k))$, $1 \leq k < n$, for high thresholds (small values of k) is observed once the diagonal section of the copula is expected to be smooth in the neighborhood of 1 and the first derivative approximately constant. However, in order to decrease variance, k cannot be too small. The algorithm proposed in Frahm *et al.* ([7], 2005) aims to identify the plateau, i.e., the stability region which is induced by the homogeneity. More precisely, first we smooth the graph $(k, \hat{\lambda}(k))$ by a box kernel with bandwidth $w = \lfloor bn \rfloor \in \mathbb{N}$ consisting of the means of $2w + 1$ successive points of $\hat{\lambda}(i)$, $i = 1, \dots, n$. Now, in the smoothed moving average values, $\bar{\lambda}(1), \dots, \bar{\lambda}(n - 2w)$, the plateaus with length $m = \lfloor \sqrt{n - 2w} \rfloor$ are defined as $p_k = (\bar{\lambda}(k), \dots, \bar{\lambda}(k + m - 1))$, $k = 1, \dots, n - 2w - m + 1$. The algorithm stops at the first plateau fulfilling the criterium

$$\sum_{i=k+1}^{k+m-1} |\bar{\lambda}(i) - \bar{\lambda}(k)| \leq 2\sigma,$$

with σ corresponding to the standard deviation of $\bar{\lambda}(1), \dots, \bar{\lambda}(n - 2w)$, and the TDC estimate corresponds to

$$\hat{\lambda} = \frac{1}{m} \sum_{i=1}^m \bar{\lambda}(k + i - 1).$$

If no plateau fulfills the stopping condition, the TDC is estimated as zero.

Observe that, if the diagonal section of the copula follows a power law, the homogeneity of $\hat{\lambda}^{\text{LOG}}$ still holds for larger k and larger bandwidths may be chosen in order to reduce the variance.

We simulate 1000 independent random samples of sizes $n = 250, 1000, 2500$, from the models:

- bivariate Normal with $\rho = 0.5$ and $\rho = 0.85$ ($\lambda = 0$; $\eta = 0.75, 0.925$, respectively);
- bivariate Student- t with $\rho = 0.5$, $\nu = 1.5$ and $\rho = 0$, $\nu = 2$ ($\lambda = 0.4406, 0.2254$, respectively; $\eta = 1$);
- Logistic with dependence parameter $r = 1/1.56$ ($\lambda = 0.4406$; $\eta = 1$) (Ledford and Tawn [12, 13], 1996/1997);
- Asymmetric Logistic with dependence parameter $r = 1/2.78$ and asymmetry parameters $t_1 = 0.5$ and $t_2 = 0.9$ ($\lambda = 0.4406$; $\eta = 1$) (Ledford and Tawn [12, 13], 1996/1997);
- Morgenstern with dependence parameter $r = 0.75$ ($\lambda = 0$; $\eta = 0.5$) (Ledford and Tawn [12, 13], 1996/1997);

We apply the algorithm described above to the tail dependence coefficients estimated by $\hat{\lambda}^{\text{SEC}}$, $\hat{\lambda}^{\text{LOG}}$, $\hat{\eta}$ and $\hat{\chi}$, defined in, respectively, (8), (9), (13) and (14), as well as, to the tail independence tests (16) and (18). In the sequel we denote (16) as test 1 and (18) as test 2. The variances within test 1 and test 2, respectively, $\hat{\sigma}_1^2 = \hat{\eta}^2/k$ and $\hat{\sigma}_2^2 = \hat{\sigma}^2/k$ with $\hat{\sigma}^2$ given in (17), are estimated by applying the algorithm to the plots $(k, \hat{\sigma}_i^2(k))$, $i = 1, 2$, but we pick the plateau at the same location of the one given by the respective coefficient estimation. In all the cases we consider the values $b = 0.0025, 0.005, 0.01, 0.015$. The boundary cases of a bivariate Normal with $\rho = 0.85$ (tail independent model but with $\eta = 0.925 \approx 1$), and a bivariate Student- t with $\rho = 0$ and $\nu = 2$ (tail dependent model with a very low TDC of 0.2254) are included in simulations in order to assess the robustness of the method.

Observe in Figures 2 and 3 that estimators $\hat{\lambda}^{\text{LOG}}$ and $\hat{\lambda}^{\text{SEC}}$ behave quite similar, although the former seems slightly better. The largest bias occurring for the smallest sample size is around 0.1 but for the largest one it is close to zero, which indicates a good performance. The exception relates to the Normal model, in particular the boundary case of $\rho = 0.85$. In the Normal model with $\rho = 0.5$, the largest bias is about 0.2. For small samples is preferably to choose bandwidths with $b = 0.005$ or $b = 0.01$. In all the other simulation results presented here, there are no significant differences between the considered bandwidths.

In what concerns estimators $\hat{\eta}$ and $\hat{\chi}$, the first one is clearly better (Figures 4 and 5). It is also robust within the boundary cases of Student- t ($\rho = 0, \nu = 2$) and Normal ($\rho = 0.85$), for large sample sizes. Observe that the bias and the root mean squared error results are very close of the ones obtained in Draisma *et al.* ([4], 2004), where k was chosen in a range where the overall performance seems best through an intensive simulation study. Estimator $\hat{\chi}$ only slightly outperforms $\hat{\eta}$ in the Normal model for $n = 250$. The proportion of samples in which tail dependence ($\eta = 1$) is rejected at a 5% significance level is plotted in Figure 6. The heuristic procedure has an overall good performance in both tests for large sample sizes. We can see that, under tail independence, test 2 outperforms test 1 as expected (see Section 2), whereas in the tail dependent case, test 1 is slightly better. However, they do not seem to be robust given the results within the above mentioned boundary cases, particularly in the Normal case.

4 An application: dependence of large losses within stock markets

We consider five years of negative daily log-returns (from 1996 to 2000) of Intel (INTC), Microsoft (MSFT) and General Electric (GE) stocks, which amounts to a sample size $n = 1262$. These data were analyzed in McNeil *et al.* ([14], 2005; Chapter 5). We aim to quantify the degree of a contagious risk of large losses within (INTC,MSFT), (INTC,GE) and (MSFT,GE), i.e., to investigate if the pairs (INTC,MSFT), (INTC,GE) and (MSFT,GE) present tail dependence or independence and quantify the respective degree of extremal dependence. As a preliminary step, we analyze the scatter plots in Figure 7. Observe that the largest values for one variable correspond to moderately large values of the same sign for the other variable, insinuating the variables are asymptotically independent but not perfectly. In Table 1 are the estimates of $\hat{\eta}$, $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\chi}$, $\hat{\lambda}^{\text{SEC}}$ and $\hat{\lambda}^{\text{LOG}}$. The results correspond to $b = 0.005$, which are very close to the ones obtained with the other bandwidths ($b = 0.0025, 0.01, 0.015$) and thus omitted. Both tests reject dependency in (INTC,MSFT) and (INTC,GE). Observe the small values provided by the TDC estimators. In the case (MSFT,GE), test 2 rejects the dependence condition and test 1 does not reject it for very little. The values of $\hat{\lambda}^{\text{SEC}}$ and $\hat{\lambda}^{\text{LOG}}$ are also small indicating that tail independence may be a more plausible conclusion. Therefore, we find that the contagious risk of large losses is residual, particularly in the case (INTC,GE).

Table 1: Estimates of $\hat{\eta}$, $\hat{\sigma}_1$, $\hat{\sigma}_2$, $\hat{\chi}$, $\hat{\lambda}^{\text{SEC}}$ and $\hat{\lambda}^{\text{LOG}}$, for (INTC,MSFT), (INTC,GE) and (MSFT,GE), with $b = 0.005$.

	$\hat{\eta}$	$\hat{\sigma}_1$	$\hat{\sigma}_2$	$\hat{\chi}$	$\hat{\lambda}^{\text{SEC}}$	$\hat{\lambda}^{\text{LOG}}$
(INTC,MSFT)	0.7321	0.0224	0.0149	0.5741	0.2629	0.2489
(INTC,GE)	0.5549	0.0065	0.0042	0.3040	0.0551	0.0372
(MSFT,GE)	0.7300	0.0321	0.0241	0.3808	0.1762	0.1613

5 Final remarks

In this paper we address the tail dependence inference problem since it is important to distinguish the type of tail dependence in order to correctly evaluate the risk of simultaneous extreme events. Most of the non parametric estimators have to deal with the choice of the number k of order statistics to consider in the production of an estimate. This is not an easy task since it requires a trade-off between variance and bias (small values of k cause large variance and large values of k increase the bias). An optimal choice of k that leads to the smallest mean squared error is difficult to derive and in practice, this is frequently solved through intensive simulation studies (see, e.g., Draisma *et al.* [4], 2004). This is also a very common problem in the estimation of the tail index, a parameter of major importance within extreme value theory (see, e.g., Beirlant *et al.* [1] 2012 and references therein). Since the non parametric estimators yield a characteristic plateau while plotting the estimates for successive k , Frahm *et al.*

([7], 2005) introduced a simple plateau-finding algorithm after smoothing the latter plot by some box kernel in order to find the optimal threshold k . Here we have applied this heuristic procedure to estimators of the TDC in (1), as well as estimators of the tail independence such as the Ledford and Tawn coefficient η in (3) and coefficient $\bar{\chi}$ in (7), for several box kernel bandwidths. We have also analyzed this methodology in two tests for tail independence given in (16) and (18). We conclude that the procedure has an overall good performance, specially for large samples. Some care must be given to the tests as they might not be robust, in particular for boundary cases within the Normal model. We call the attention for the very good performance of η estimation. We recall that it is based on a tail index estimator (Hill estimator) which may be an indication that this procedure can also work well within the tail index estimation. This will be addressed in a future work.

References

- [1] Beirlant, J., Caeiro, F., Gomes, M.I. (2012). An overview and open research topics in statistics of univariate extremes. *RevStat* 10(1), 1–31.
- [2] Beirlant, J., Goegebeur, Y., Segers, J., Teugels, J. (2004). *Statistics of Extremes: Theory and Application*. New York: Wiley.
- [3] Coles, S., Heffernan, J., Tawn, J. (1999). Dependence measures for extreme value analysis. *Extremes* 2, 339–366.
- [4] Draisma, G., Drees, H., Ferreira, A., de Haan, L. (2004). Bivariate tail estimation: dependence in asymptotic independence. *Bernoulli* 10, 251–280.
- [5] Embrechts, P., McNeil, A., Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In: *Risk Management: Value at Risk and Beyond*, ed. M.A.H. Dempster, Cambridge University Press, Cambridge, 176–223.
- [6] Ferreira, M. (2013). Nonparametric estimation of the tail-dependence coefficient. *RevStat* 11(1), 1–16.
- [7] Frahm, G., Junker, M., Schmidt R. (2005). Estimating the tail-dependence coefficient: properties and pitfalls. *Insurance: Mathematics & Economics* 37(1), 80–100.
- [8] Hill, B.M. (1975). A Simple General Approach to Inference About the Tail of a Distribution. *Ann. Stat.* 3, 1163–1174.
- [9] Huang, X. (1992). *Statistics of Bivariate Extreme Values*. Ph. D. thesis, Tinbergen Institute Research Series 22, Erasmus University Rotterdam.
- [10] Joe, H. (1997). *Multivariate Models and Dependence Concepts*. Chapman and Hall, London.
- [11] Joe, H., Smith, R.L., Weissman, I. (1992). Bivariate threshold models for extremes. *Journal of the Royal Statistical Society, Series B* 54, 171–183.

- [12] Ledford, A., Tawn, J.A. (1996). Statistics for near independence in multivariate extreme values. *Biometrika* 83, 169–187.

- [13] Ledford, A. Tawn, J.A. (1997). Modelling dependence within joint tail regions. *Journal of the Royal Statistical Society. Series B* 59, 475–499.

- [14] McNeil, A.J., Frey R., Embrechts, P. (2005). *Quantitative risk management*. Princeton University Press, Princeton Series in Finance, New Jersey.

- [15] Peng, L. (1999). Estimation of the coefficient of tail dependence in bivariate extremes. *Statist. Probab. Lett.* 43, 399–409.

- [16] Poon, S.-H., Rockinger, M., Tawn, J. (2004). Extreme value dependence in financial markets: diagnostics, models, and financial implications. *Review of Financial Studies* 17 (2), 581–610.

- [17] Schmidt, R., Stadtmüller, U. (2006). Nonparametric estimation of tail dependence. *Scandinavian Journal of Statistics* 33, 307–335.

- [18] Sibuya, M. (1960). Bivariate extreme statistics. *Annals of the Institute of Statistical Mathematics* 11, 195–210.

- [19] Sklar, A. (1959). Fonctions de répartition à n dimensions et leurs marges. *Publications de l'Institut de statistique de l'Université de Paris* 8, 229–231.

The authors declare that there is no conflict of interests regarding the publication of this article.

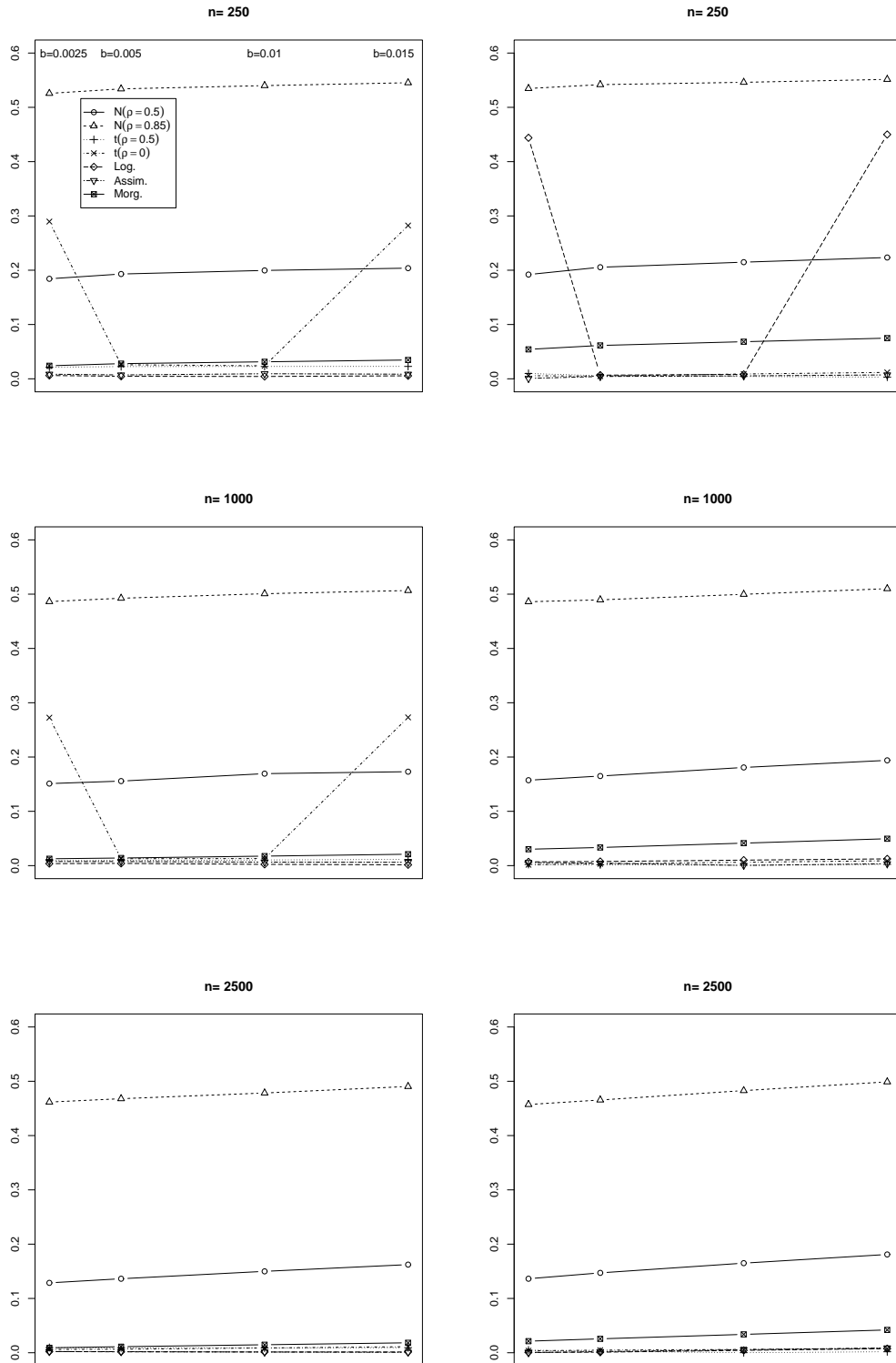


Figure 2: Absolute bias of estimator $\hat{\lambda}^{\text{LOG}}$ (first column) and estimator $\hat{\lambda}^{\text{SEC}}$ (second column). The four values plotted in each line correspond to $b = 0.0025, 0.005, 0.01, 0.015$, respectively.

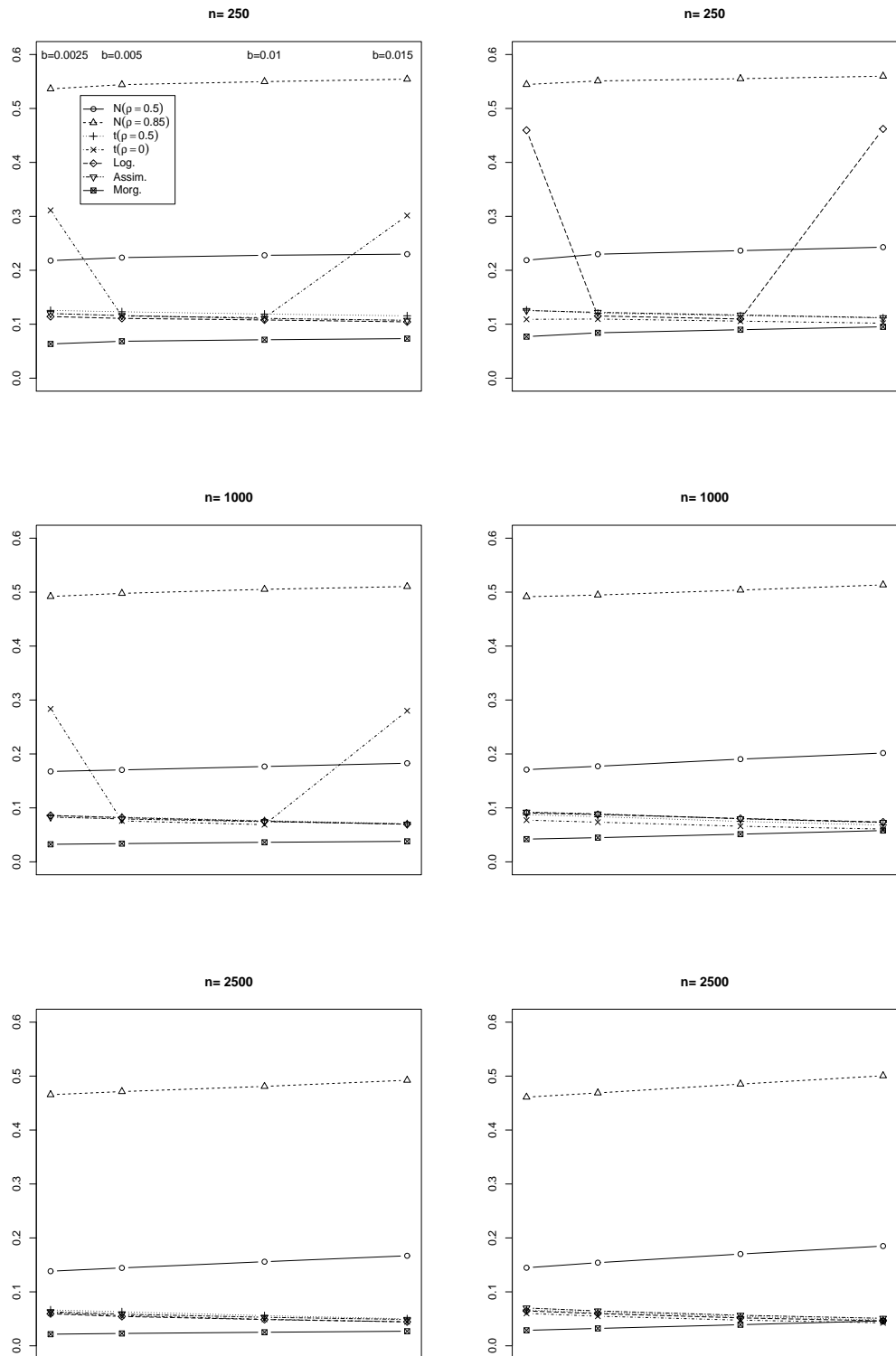


Figure 3: Root mean squared error (rmse) of estimator $\hat{\lambda}^{\text{LOG}}$ (first column) and estimator $\hat{\lambda}^{\text{SEC}}$ (second column). The four values plotted in each line correspond to $b = 0.0025, 0.005, 0.01, 0.015$, respectively.

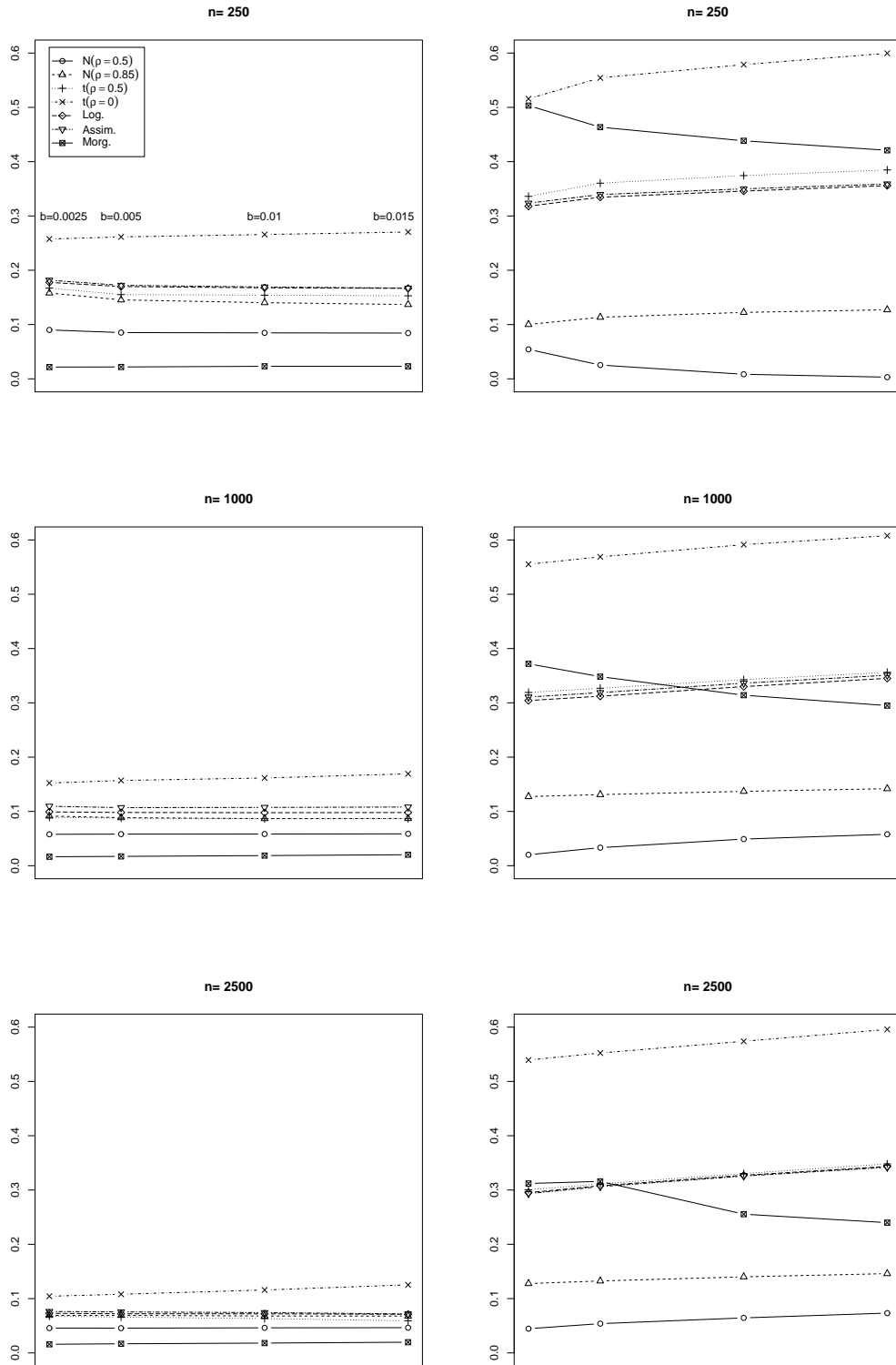


Figure 4: Absolute bias of estimator $\hat{\eta}$ (first column) and estimator $\hat{\chi}$ (second column). The four values plotted in each line correspond to $b = 0.0025, 0.005, 0.01, 0.015$, respectively.

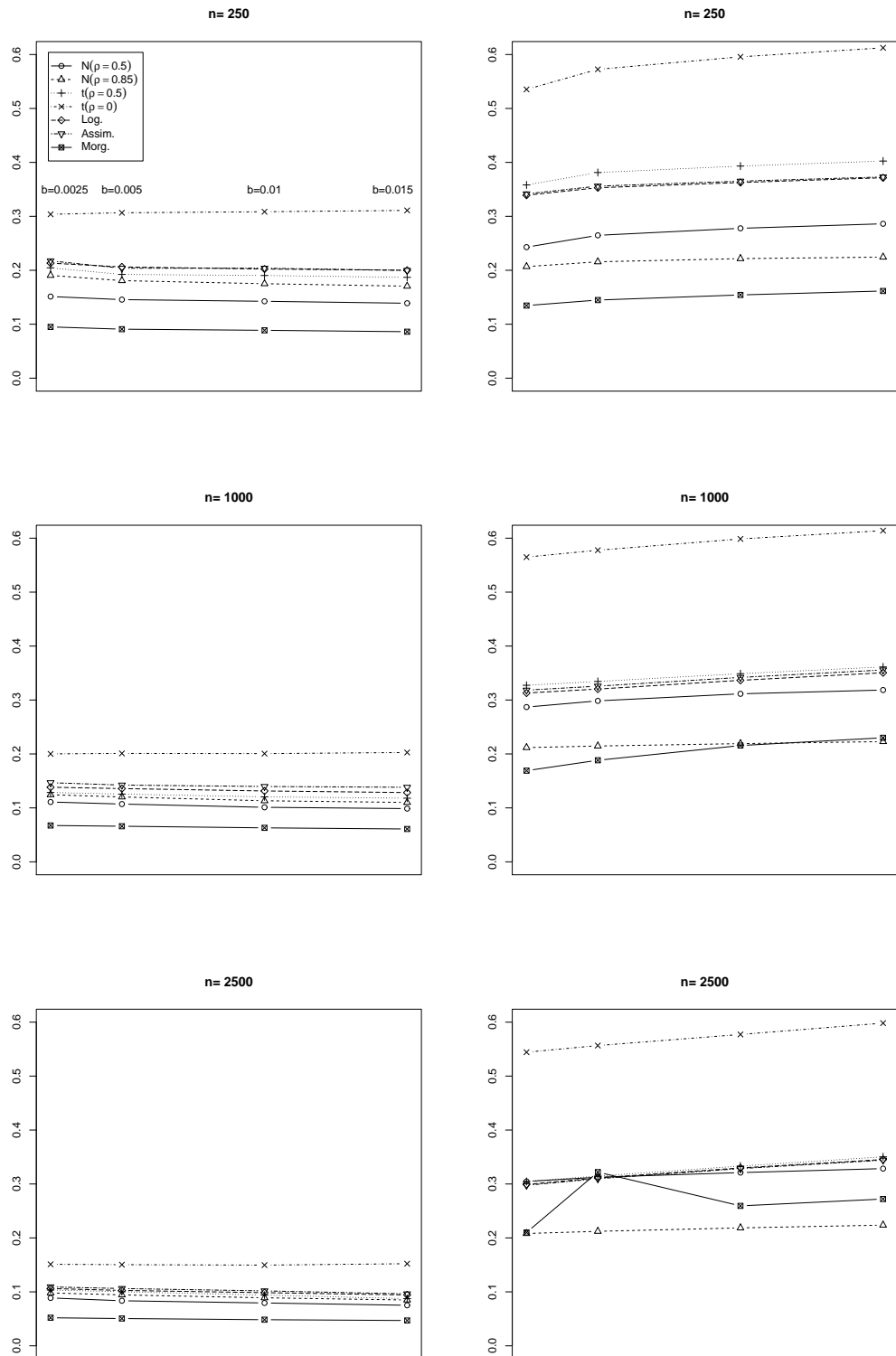


Figure 5: Root mean squared error (rmse) of estimator $\hat{\eta}$ (first column) and estimator $\hat{\chi}$ (second column). The four values plotted in each line correspond to $b = 0.0025, 0.005, 0.01, 0.015$, respectively.

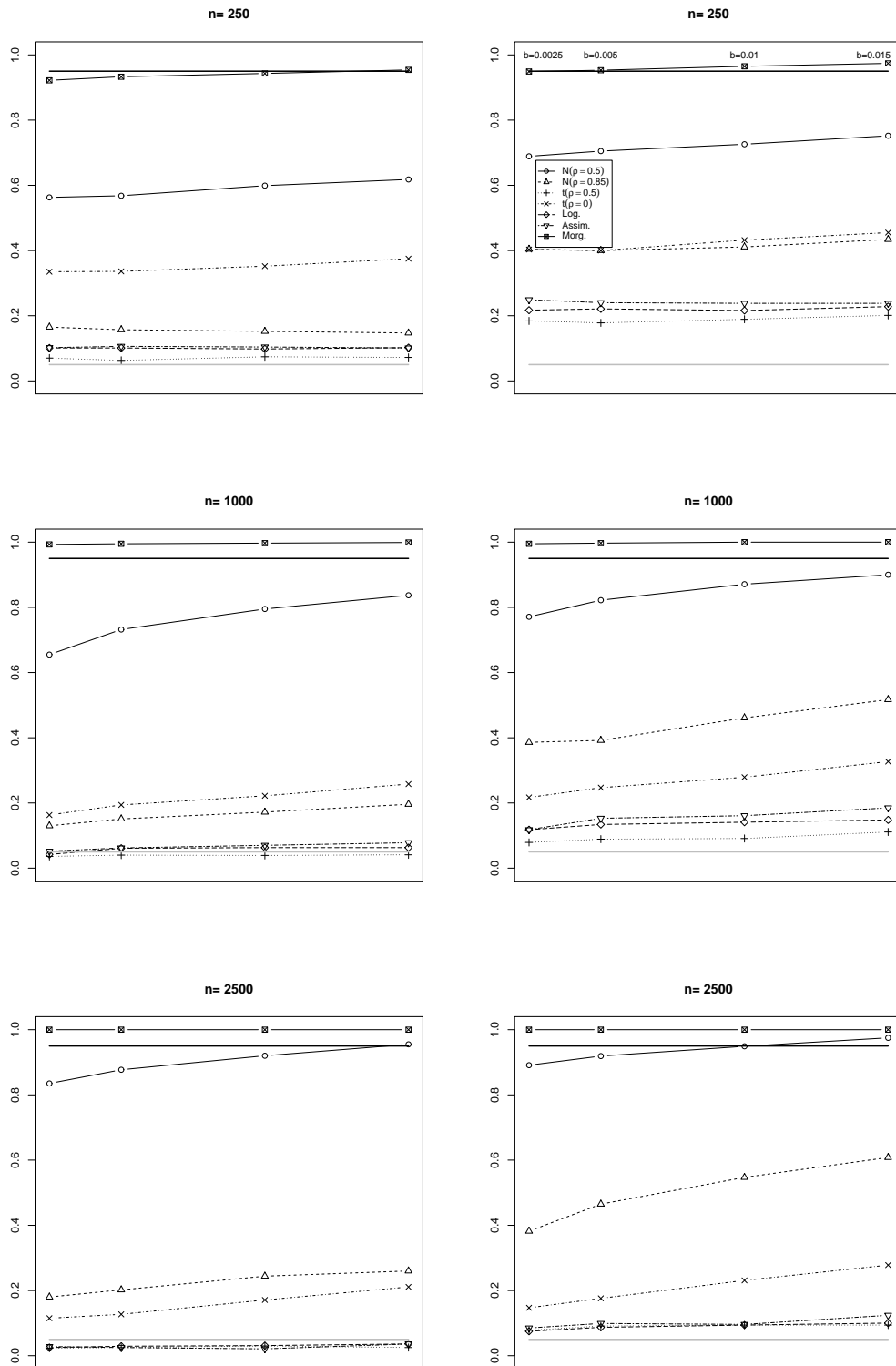


Figure 6: Proportion of samples in which $\eta = 1$ is rejected by a 5% test for test 1 (first column) and for test 2 (second column). The horizontal solid lines correspond to 95% confidence (black) and 5% significance (grey). The four values plotted in each line correspond to $b = 0.0025, 0.005, 0.01, 0.015$, respectively.

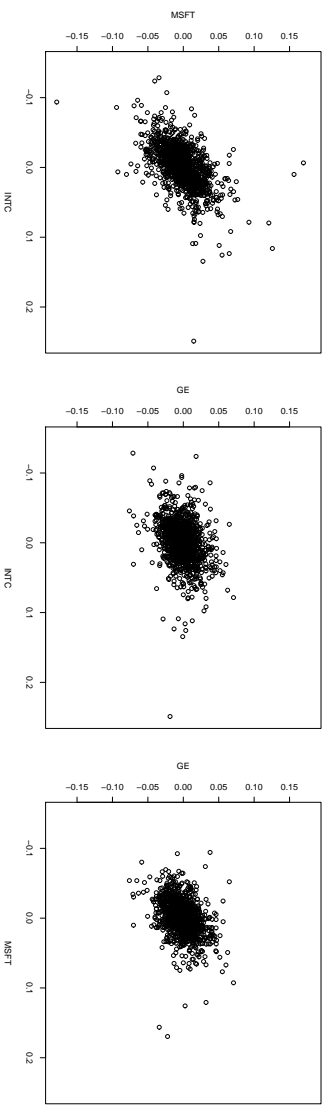


Figure 7: Left-to-right: Scatter plots of $(INTC, MSFT)$, $(INTC, GE)$ and $(MSFT, GE)$, respectively.