A. TESTING FOR PQD

A. I. Motivation

Risk management in Finance and Insurance

A. II. Stochastic orderings

Review of concepts

A. III. Sklar's representation

Dependence structure described by copulas

A.IV. Dependence notions

A.IV.1. Positive quadrant dependence

Probability that two random variables are small is at least as great as it would be were they independent

A.IV.2. Positive orthant dependence

Extension in higher dimensions

A.V. Applications of dependence notions

A.V.1. Positive quadrant dependence

- Implication for
  pricing insurance premiums and options

- Link with Lorenz order
  (inequality measurement)

A.V.2. Positive lower orthant dependence

- Implication for insurance premiums

A.VI. Hypotheses testing

A.VI.1. Inference based on cdfs

Asymptotic normality of estimator
(loss level)

A.VI.2. Inference based on copulas

Asymptotic normality of estimator
(probability level)

A.VI.3. Testing procedures

- Distance test
for inequality constraints

- Intersection-union test
for inequality constraints

A.VII. Empirical illustrations

A.VII.1. US losses and ALAE's

- PQD between losses
  and allocated loss adjustment expenses

- Comparison
  between reinsurance premiums
  under different dependence assumptions

A.VII.2. Danish fire losses

- PQD between damage to buildings
  and damage to furniture

- Comparison between stop-loss premiums
  under different dependence assumptions

A.VII.3. Hedge fund and stock indices

A.VIII. Conclusions

A.I. Motivation

Why testing
for positive quadrant dependence?

Risk Management in Finance and Insurance:

Pressure from
Regulatory Environment
/ Internal Risk Managers

⇒ Development of
Proprietary Risk Measurement Models

Dependence structure between asset prices affects risk measures and portfolio selection

Main risk in large portfolios is occurrence of many joint default events or simultaneous downside evolution of prices

Positive Quadrant Dependence (PQD)
describes clustering behavior

Definition :

PQD if probability that two random variables are small is at least as great as it would be were they independent

PQD also allows risk manager to compare the sum of PQD with corresponding sum under independence assumption

Comparison is
in the sense of stochastic orderings

Thus

Need of testing procedures
for presence of PQD

A.II.. Stochastic orderings

U_1 = class of all non-decreasing utilities

U_2 = restriction of U_1 to its concave elements

Remarks:

$u' \geq 0$: increasing (prefer more than less)

$u'' \leq 0$: concave (risk aversion)

Y_1 is said to be smaller than Y_2

- in the stochastic dominance order

  if $Eu(Y_1) \leq Eu(Y_2)$ for $u \in U_1$

denoted as $Y_1 \prec_d Y_2$

- in the increasing concave order

  if $Eu(Y_1) \leq Eu(Y_2)$ for $u \in U_2$

denoted as $Y_1 \prec_{icv} Y_2$

- Y_2 is preferred over Y_1 by all risk averters

  if $Y_1 \prec_{icv} Y_2$ and $EY_1 = EY_2$

denoted as $Y_1 \prec_{cv} Y_2$

Main characterizations:

(i) $Y_1 \prec_d Y_2 \Leftrightarrow F_1(x) \geq F_2(x), \forall x$

(ii) $Y_1 \prec_d Y_2 \Leftrightarrow$

$$F_1^{-1}(p) \leq F_2^{-1}(p), \forall p \in (0,1)$$

(iii) $Y_1 \prec_{icv} Y_2 \Leftrightarrow$

$$\int_{-\infty}^{x} F_1(u)du \geq \int_{-\infty}^{x} F_2(u)du, \forall x$$

(iv) $Y_1 \prec_{icv} Y_2 \Leftrightarrow$

$$\int_0^p F_1^{-1}(u)du \leq \int_0^p F_2^{-1}(u)du, \forall p \in (0,1)$$

$\Rightarrow$ Statistical inference based on empirical distributions evaluated at predetermined grid of d points and set of inequalities $D^i \geq 0, i = 1,\ldots,d$.

A.III. Sklar's representation

Data :

$$\{Y_t = (Y_{1t},\ldots,Y_{nt})', t = 1,\ldots,T\}$$

= i.i.d. observations
(observed returns or losses)

Distributions :

$f(y), F(y)$ = joint pdf and cdf of Y_t

$f_j(y_j), F_j(y_j)$ = pdf and cdf of margins

$\Rightarrow$ copula describes how joint distribution F is ``coupled'' to its univariate margins F_j

copula = summary of dependence structure

Sklar's Theorem :

Let F be an n-dimensional distribution function with margins $F_1,\ldots,F_n$. Then there exists an n-copula C:

$$F(y_1,\ldots,y_n) = C(F_1(y_1),\ldots,F_n(y_n)) \quad (1)$$

If $F_1,\ldots,F_n$ are all continuous,
then C is uniquely defined.

Conversely, if C is an n-copula and $F_1,\ldots,F_n$ are distribution functions, then the function F defined by (1) is an n-dimensional distribution function with margins $F_1,\ldots,F_n$

Corollary:

$$C(u_1,...,u_n) = F(F_1^{-1}(u_1),...,F_n^{-1}(u_n))$$

⇒ copulas are multivariate uniform distributions which describe the dependence structure of random variables

Main properties :

1. Strictly increasing transformations of random variables result in transformed variables having the same copula (not true for correlation)

2. Independence ⇔ $C(u) = \prod_{j=1}^{n} u_j$

Comonotonicity ⇔ $C(u) = \min(u_1,...,u_n)$ (each variable is almost surely a strictly increasing function of any of the others)

3. Links with dependence measures

Kendall's tau:

$$\tau_{Y_1,Y_2} = 1 - 4\int_0^1\int_0^1 \frac{\partial C(u_1,u_2)}{\partial u_1}\frac{\partial C(u_1,u_2)}{\partial u_2} du_1 du_2$$

(association measure on ranks)

Spearman's rho:

$$\rho_{Y_1,Y_2} = 12\int_0^1\int_0^1 C(u_1,u_2) du_1 du_2 - 3$$

(correlation on the ranks instead of observations)

A.IV. Dependence notions

A.IV.1. Positive quadrant dependence

Y_1 and Y_2 are PQD if

$$P\lfloor Y_1 \leq y_1, Y_2 \leq y_2 \rfloor \geq P\lfloor Y_1 \leq y_1 \rfloor P\lfloor Y_2 \leq y_2 \rfloor$$

PQD if probability that two random variables are small is at least as great as it would be were they independent

In terms of copulas:

$$C(u_1, u_2) \geq u_1 u_2$$

A.IV.2 Positive orthant dependence

$(Y_1,\ldots,Y_n)'$ are PLOD if

$$P\lfloor Y_1 \le y_1,\ldots,Y_n \le y_n \rfloor \ge P\lfloor Y_1 \le y_1 \rfloor .. P\lfloor Y_n \le y_n \rfloor$$

(Positive Lower Orthant Dependent)

In terms of copulas:

$$C(u_1,\ldots,u_n) \ge u_1 \ldots u_n$$

PUOD (Positive Upper Orthant Dependent)
if > substituted for ≤

A.V. Applications of dependence notions

A.V.1. Positive quadrant dependence

Result: If Y_1 and Y_2 are PQD,

$$\text{then } Y_1 + Y_2 \prec_{cv} Y_1^{\perp} + Y_2^{\perp}$$

- Implication for pricing insurance premiums and options

When PQD holds, every risk-averter agrees to say that $Y_1 + Y_2$ is less favorable than the corresponding sum under independence, namely $Y_1^{\perp} + Y_2^{\perp}$

Most insurance premiums and risk measures will be larger for $Y_1 + Y_2$ than for $Y_1^{\perp} + Y_2^{\perp}$

Example:

$$E\left[\left(Y_1^{\perp} + Y_2^{\perp} - K\right)_+\right] \leq E\left[\left(Y_1 + Y_2 - K\right)_+\right]$$

= stop-loss premium in actuarial science

= price of a basket option in finance

- Link with Lorenz order
  (inequality measurement)

Lorenz order defined by
pointwise comparison of Lorenz curves:

$$L(p) = \frac{1}{EY}\int_0^p F^{-1}(u)du, \forall p \in (0,1)$$

When Y = income of individuals, L maps $p \in (0,1)$ to proportion of total income of population which accrues to poorest 100p% of population

Y_1 is said to be smaller than Y_2 in the Lorenz order, when

$$L_1(p) \geq L_2(p), \forall p \in (0,1)$$

denoted by $Y_1 \prec_{Lorenz} Y_2$

If $EY_1 = EY_2$, $Y_1 \prec_{Lorenz} Y_2 \Leftrightarrow Y_2 \prec_{cv} Y_1$

If Y_1 and Y_2 are PQD,

then $Y_1^{\perp} + Y_2^{\perp} \prec_{Lorenz} Y_1 + Y_2$

Interpretation :

Y_1, resp. Y_2, = husbands', resp. wives', income

Y_1 and Y_2 are PQD means ``birds of a feather flock together''

men and women earning large (resp. small) salaries tend to be associated

Population exhibits more inequality in the Lorenz sense than a population where spouses' earnings are independent

A.V.2. Positive lower orthant dependence

If remaining lifetimes are PLOD,
net single premium of a joint life annuity, which pays USD 1 each year if all individuals survive is higher than under independence

A.VI. Hypotheses testing

A.VI.1. Inference based on cdfs

Take *d* points $y_i = (y_{i1},\ldots,y_{in})'$

and define $$D_F^{\,i} = F(y_i) - \prod_{j=1}^{n} F_j(y_{ij})$$

and $D_F = (D_F^{\,1},\ldots,D_F^{\,d})'$

- Null hypothesis of a test for PLOD:

$$H_F^{\,0} = \{D_F : D_F \geq 0\}$$

Alternative hypothesis:

$$H_F^{\,1} = \{D_F : D_F \text{ unrestricted}\}$$

- Null hypothesis of a test for *non*-PLOD:

$$\overline{H}_F^{\,0} = \{D_F : D_F^{\,i} \leq 0 \text{ for some } i\}$$

Alternative hypothesis:

$$\overline{H}_F^{\,1} = \{D_F : D_F^{\,i} > 0 \text{ for all } i\}$$

- Empirical counterparts $\hat{D}_F{}^i$ of $D_F{}^i$ obtained by substituting empirical distributions for unknown distributions:

$$\hat{F}(y_i) = \frac{1}{T}\sum_{t=1}^{T}\prod_{j=1}^{n} I\left[Y_{jt} \le y_{ij}\right], i = 1,\ldots,d$$

$$\hat{F}_j(y_{ij}) = \frac{1}{T}\sum_{t=1}^{T} I\left[Y_{jt} \le y_{ij}\right], i = 1,\ldots,d$$

Remark :

Asymptotic normality of $\sqrt{T}(\hat{D}_F - D_F)$ can be deduced from the asymptotic normality of the above quantities.

A.VI..2. Inference based on copulas

Take d points $u_i = (u_{i1},\ldots,u_{in})'$,
with $u_{ij} \in (0,1)$

and define $$D_C{}^i = C(u_i) - \prod_{j=1}^{n} u_{ij}$$

and $D_C = (D_C{}^1,\ldots,D_C{}^d)'$

- Null hypothesis of a test for PLOD:

$$H_C{}^0 = \{D_C : D_C \geq 0\}$$

Alternative hypothesis:

$$H_C{}^1 = \{D_C : D_C \text{ unrestricted}\}$$

- Null hypothesis of a test for *non*-PLOD:

$$\overline{H}_C^{\,0} = \left\{ D_C : D_C^{\,i} \leq 0 \ \text{for some}\ i \right\}$$

Alternative hypothesis:

$$\overline{H}_C^{\,1} = \left\{ D_C : D_C^{\,i} > 0 \ \text{for all}\ i \right\}$$

- Empirical counterpart $\hat{D}_C^{\,i}$ of $D_C^{\,i}$ obtained by $\hat{C}(u_i) = \hat{F}(\hat{\varsigma}_i)$

where $\hat{\varsigma}_i = (\hat{\varsigma}_{i1}, \ldots, \hat{\varsigma}_{in})$ made of empirical univariate quantiles $\hat{\varsigma}_{ij}$

Levels are no more given deterministic values, but quantiles, and thus random quantities which affect the asymptotic distribution.

A.VI.3. Testing procedures

- PLOD : use of distance tests

Let $\widetilde{D}_K$, $K = F, C$, be solution of:

$$\inf_D T\left(D - \hat{D}_K\right)' \hat{V}_K \left(D - \hat{D}_K\right) \text{s.t. } D \geq 0$$

where $\hat{V}_K$ is a consistent estimate of the asymptotic variance V_K

Under the null the test statistic

$$\hat{\xi}_K = T\left(\widetilde{D} - \hat{D}_K\right)' \hat{V}_K \left(\widetilde{D} - \hat{D}_K\right)$$

should be close to zero.

The asymptotic distribution under the null is given by

$$P\left[\hat{\xi}_K \geq x\right] = \sum_{i=1}^{d} P\left[\chi_i^{\,2} \geq x\right] w(d, d-i, \hat{V}_K)$$

where the weight $w(d, d-i, \hat{V}_K)$ corresponds to the probability that $\widetilde{D}_K$ has exactly *d-i* positive elements.

* Computation of solution $\widetilde{D}_K$ performed by numerical optimization routine for constrained quadratic programming problems

* Computations of weights by Monte Carlo or use tabulated bounds

- non-PLOD : use of intersection-union tests

Let $\hat{\gamma}_K^{\,i} = \sqrt{T}\hat{D}_K^{\,i} / \sqrt{\hat{V}_{K,ii}}$, $K = F, C$

Reject $\overline{H}_K^{\,0}$ when $\inf \hat{\gamma}_K^{\,i} > z_{1-\alpha}$
where $z_{1-\alpha}$ is the $1-\alpha$ quantile of a *N(0,1)*

Power issues in intersection-union tests:

1) do not exploit covariance structure
2) highly conservative if $\hat{D}_K \cong 0$ for some coordinates

A.VII. Empirical Illustrations

Detection of PQD in US and Danish claim data

A.VII.1. US losses and ALAE's

1,466 values of pair (LOSS,ALAE) collected by US Insurance Services Office

Pearson's $r = 0.3805$ (linear correlation)

Kendall's $\tau = 0.3067$

Spearman's $\rho = 0.4437$

Remark: Work with log data since (LOSS,ALAE) PQD
$\Leftrightarrow$ (log(LOSS),log(ALAE)) PQD

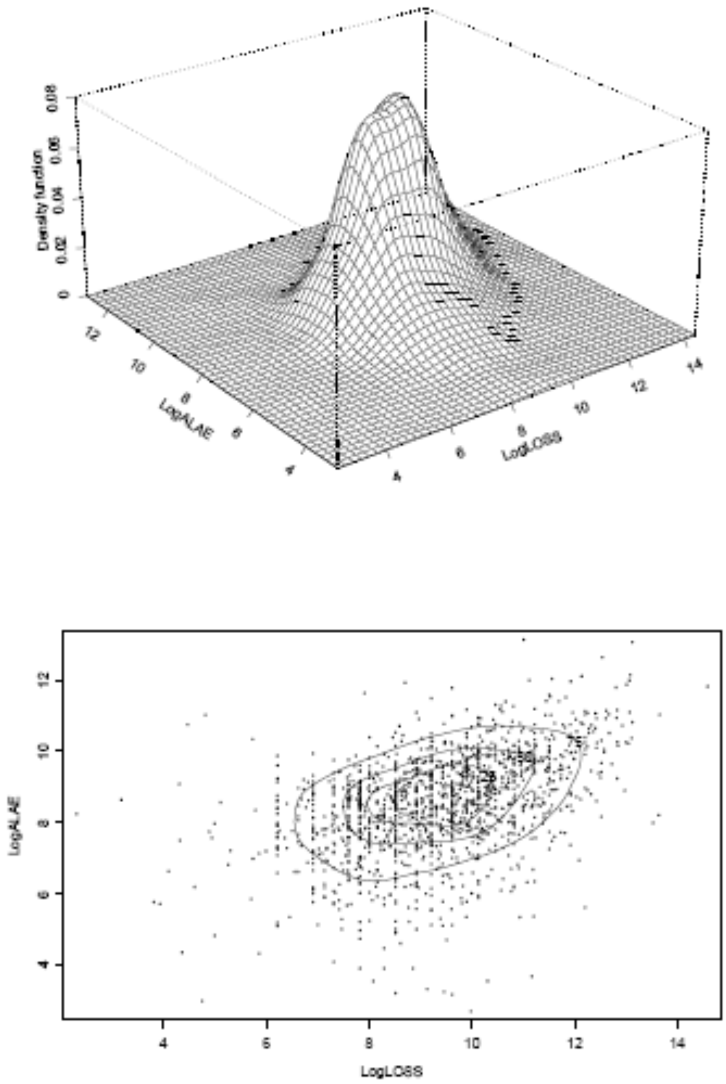

Figure 7.7.1: Kernel estimation of the bivariate pdf for (log(LOSS),log(ALAE))

	LOSS	ALAE
Mean	37,109.58	12,017.47
Std Dev.	92,512.80	26,712.35
Skew.	10.95	10.07
Kurt.	209.62	152.39
Min	10.00	15.00
Max	2,173,595.00	501,863.00
1st Quart.	3,750.00	2,318.25
Median	11,048.50	5,420.50
3rd Quart.	32,000.00	12,292.00

Table 7.1: Summary statistics for variables LOSS and ALAE

- *Tests based on distribution functions*

Whole domain:

49 points coming from the equally spaced grid $\{6,7,\ldots,12\}\times\{6,7,\ldots,12\}$

Upper tails:

49 points coming from grid
$\{10,10.3,\ldots,12\}\times\{10,10.3,\ldots.,12\}$

Distance tests: PQD not rejected

Intersection-union tests: non-PQD not rejected

- *Tests based on copulas*

Whole domain

81 deciles coming from the equally spaced grid $\{.1,.2,\ldots,.9\}\times\{.1,.2,\ldots,.9\}$

Upper tails

91 percentiles coming from grid $\{.91,.92,\ldots,.99\}\times\{.91,.92,\ldots,.99\}$

Distance tests: PQD not rejected

Intersection-union tests: non-PQD not rejected

- Implication for reinsurance premiums

Reinsurance treaty on a policy with unlimited liability and insurer's retention R

Reinsurer's payment:

$$g(LOSS, ALAE)$$

$$= \begin{cases} 0 \text{ if } LOSS \leq R \\ LOSS - R + \dfrac{LOSS - R}{LOSS} ALAE \text{ if } LOSS > R \end{cases}$$

Pure premium:

$$\pi = Eg(LOSS, ALAE)$$

Estimation under:

1. independence, i.e.

$$\hat{\pi} = \frac{1}{T^2} \sum_{t=1}^{T} \sum_{t'=1}^{T} g(LOSS_t, ALAE_t)$$

2. dependence expressed by the data, i.e.

$$\hat{\pi} = \frac{1}{T} \sum_{t=1}^{T} g(LOSS_t, ALAE_t)$$

3. classical comonotonic approximation for (LOSS,ALAE), i.e.

$$\hat{\pi} = \frac{1}{T} \sum_{t=1}^{T} g(LOSS_t, \hat{F}_2^{-1}(\hat{F}_1(LOSS_t)))$$

R	10,000	50,000	100,000	500,000	1,000,000
indep.	33,308.9054	19,108.3604	12,402.7515	1,800.9984	804.9684
dep.	36,765.8687	21,227.8071	13,801.1927	1,875.0277	850.1686
comon.	38,962.6734	23,271.1908	15,407.7782	2,308.0139	985.3801

Table 5.3: Pure premiums for a reinsurance treaty with retention R.

A.VII.2. Danish fire losses

1,485 values of pair (BUILDINGS/CONTENTS) for 1980 to 1990

damage to buildings / damage to contents (furniture)

Pearson's r = 0.5362

Kendall's τ = 0.0741

Spearman's ρ = 0.1385

Remark: Work with log data

	Buildings	Contents
Mean	1,731,012	1,391,979
Std Dev.	2,842,519	3,776,137
Skew.	12.102	9.068
Kurt.	217.587	122.316
Min	25,000	10,000
Max	65,000,000	72,500,000
1st Quart.	800,000	250,000
Median	1,100,000	430,000
3rd Quart.	1,775,000	1,000,000

Table 7.4: Summary statistics for variables Buildings and Contents.

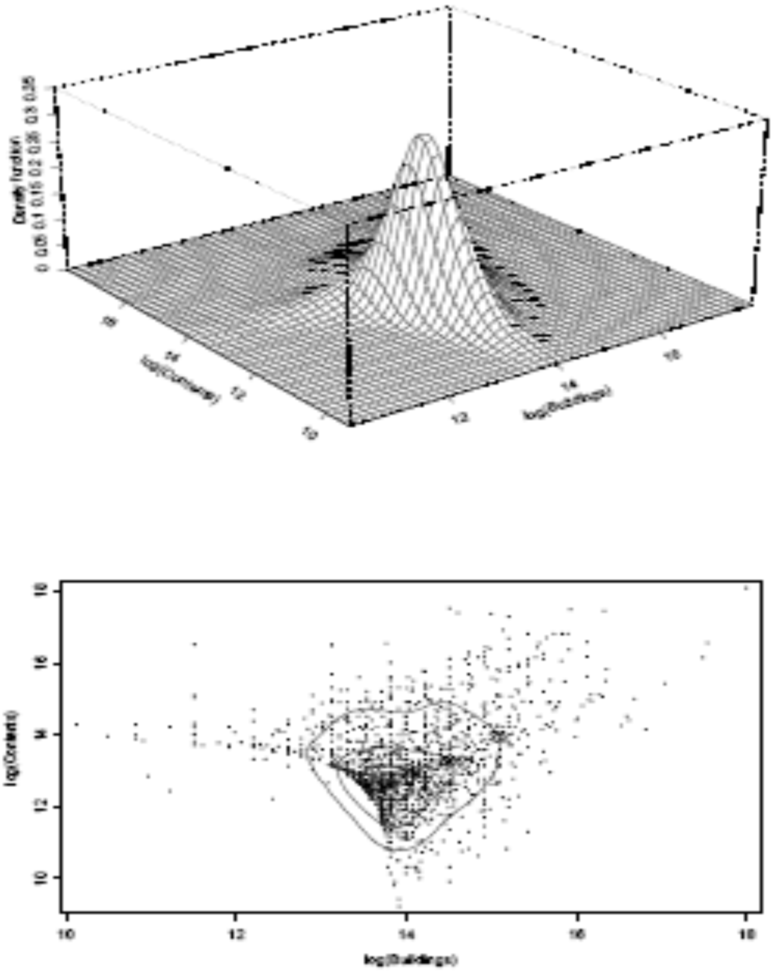

Figure 7.7.2: Kernel estimation of the bivariate pdf for (log(Buildings),log(Contents)).

- *Tests based on distribution functions*

Whole domain :

49 points coming from the equally spaced grid $\{11,12,\ldots,17\}\times\{11,12,\ldots,17\}$

Upper tails:

49 points coming from grid $\{15,15.3,\ldots,17\}\times\{15,15.3,\ldots,17\}$

Distance tests: PQD not rejected

Intersection-union tests: non-PQD not rejected

- *Tests based on copulas*

Whole domain

81 deciles coming from the equally spaced grid $\{.1,.2,\ldots,.9\}\times\{.1,.2,\ldots,.9\}$

Upper tails

91 percentiles coming from grid

$$\{.91,.92,\ldots,.99\} \times \{.91,.92,\ldots,.99\}$$

Distance tests: PQD not rejected

Intersection-union tests: non-PQD not rejected

- *Implication for stop-loss premiums*

Stop-loss premiums with deductible κ :

$$E(Buildings + Contents - \kappa)_+$$

Estimation under independence, data dependence, comonotonicity

κ ($\times 10^5$)	5	7.5	10	12.5	15	17.5	20
indep.	782,667	556,565	420,241	328,537	265,217	219,057	185,606
dep.	919,680	698,131	550,809	439,888	356,078	297,493	254,729
comon.	1,046,330	802,787	652,342	545,302	461,813	397,181	348,361

Table 7.5: Stop-loss premiums for different deductibles κ.

A.VII.3. Hedge fund and stock indices

118 values of monthly returns on pairs (S&P500/HFR) and (S&P500/CSFB-Tremont) for 31/01/1994 to 31/10/2003

Market neutral hedge fund indices HFR and CSFB-Tremont

(S&P500/HFR)

Pearson's r = 0.1406
Kendall's τ = 0.0630
Spearman's ρ = 0.0875

(S&P500/CSFB-Tremont)

Pearson's r = 0.3952
Kendall's τ = 0.2543
Spearman's ρ = 0.3728

HFR	
Mean	0.0070372
Std Dev.	0.0094661
Skew.	0.2279256
Kurt.	0.4175570
Min	-0.0167000
Max	0.0359000
1st Quart.	0.0015250
Median	0.0063000
3rd Quart.	0.01.2975
CSFB/Tremont	
Mean	0.0085466
Std Dev.	0.0089309
Skew.	0.2007868
Kurt.	0.1991623
Min	-0.0115000
Max	0.0032600
1st Quart.	0.0028250
Median	0.0081000
3rd Quart.	0.0142750
S&P500	
Mean	0.0068819
Std Dev.	0.0462238
Skew.	-0.7155344
Kurt.	0.5972095
Min	-0.1575860
Max	0.0923238
1st Quart.	-0.0220540
Median	0.0122834
3rd Quart.	0.0391559

Table 5.4: Summary statistics for HFR, CSFB/Tremont and S&P500.

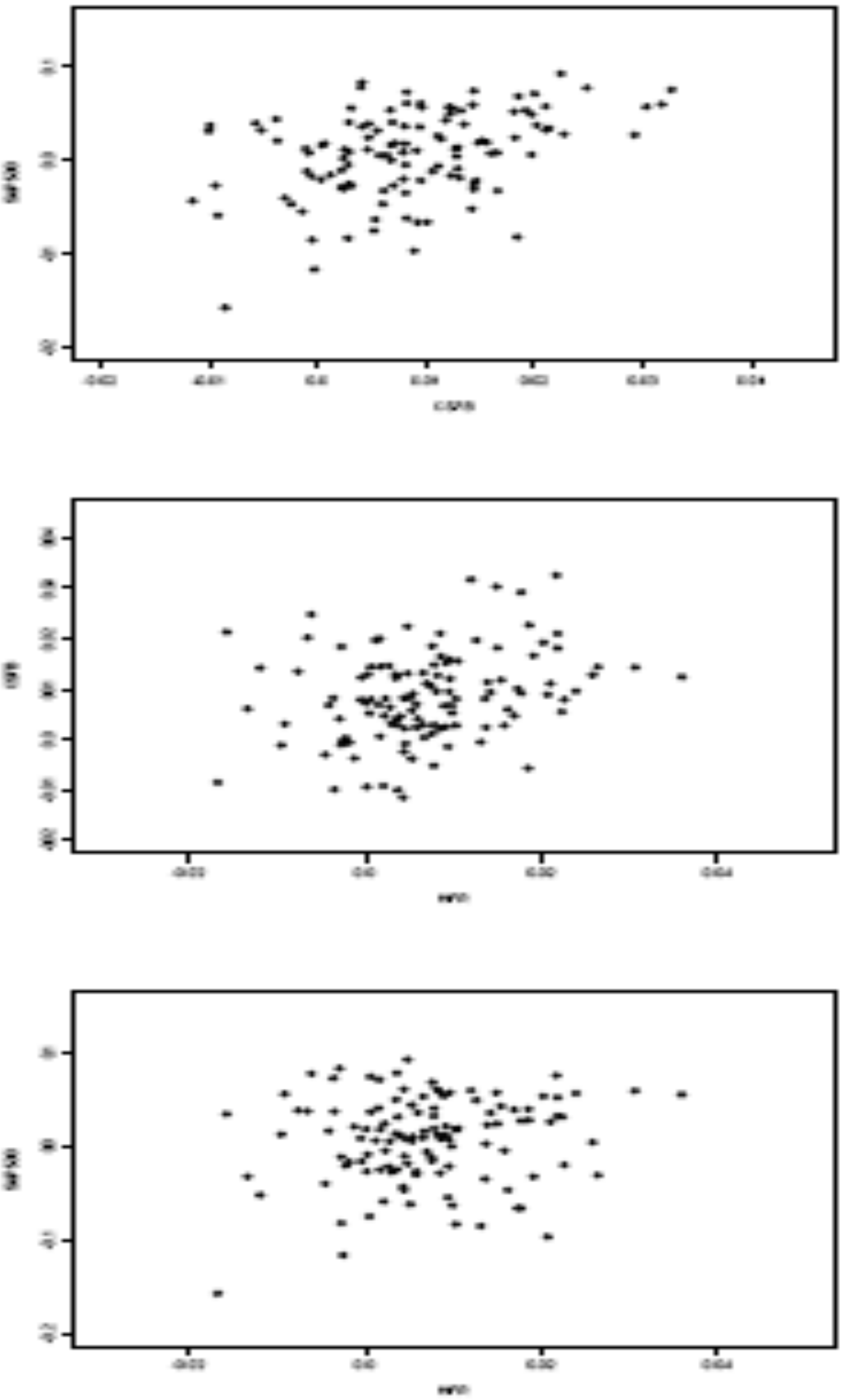

Figure 5.2: Bivariate scatterplots of the data.

- *Tests based on copulas*

cf. regulators work with probability levels.

Whole domain

16 quantiles coming from the equally spaced grid $\{.2,.4,...,.8\}\times\{.2,.4,...,.8\}$

Distance tests: PQD not rejected

Intersection-union tests: non-PQD not rejected

A.VIII. Conclusions

1) Simple distributional free inference for PQD and PLOD

2) Empirically relevant to analysis of dependencies among claim data

Show Strong PQD nature of claim data and indices

Help to achieve better design of insurance contracts in terms of premium valuation

It is also possible to design testing procedures based on other principles, such as Kolmogorov-Smirnov type of tests for PQD