Multivariate Copulas

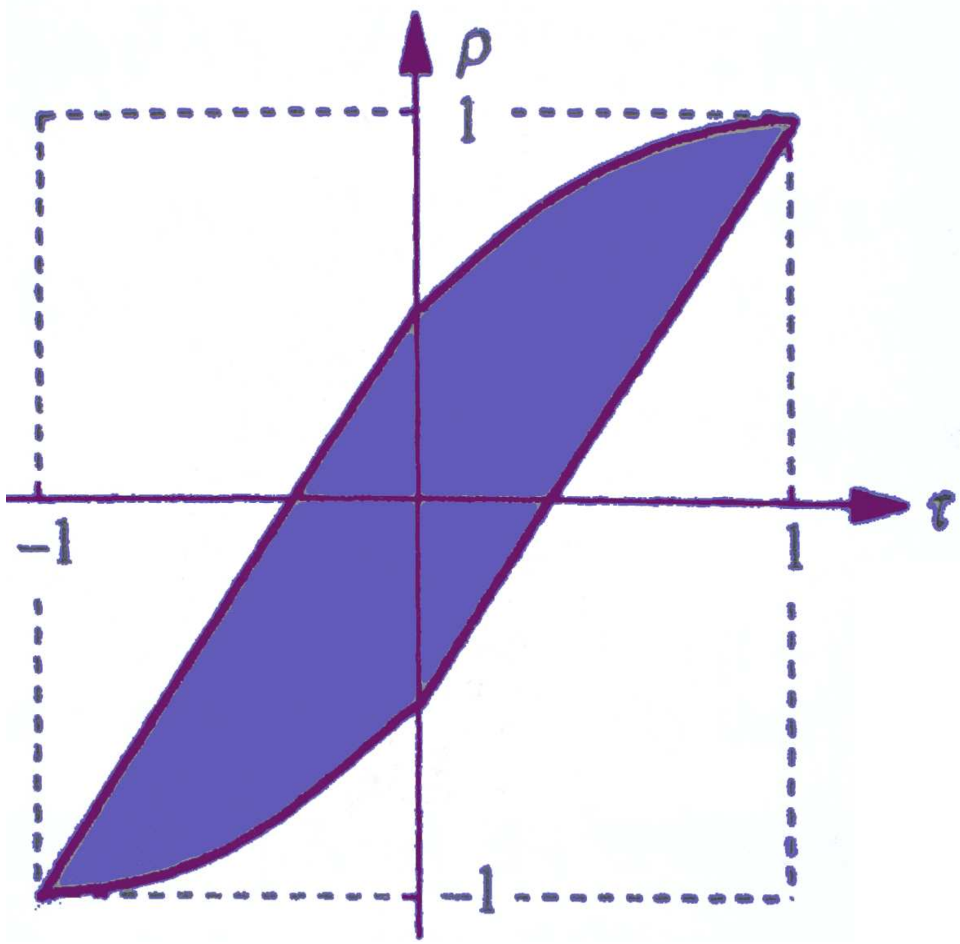
Gary Venter

CAS Spring Meeting '008

Correlation

- Kendall tau and rank correlation depend only on copula, not marginals
  - Not true for linear correlation
- Can calculate tau as:
  - $4E(C(u,v)) - 1 = (\iint_{I^2} C(u,v)c(u,v)\,du\,dv - \frac{1}{4})/(1/4)$
  - and rho as $(\iint_{I^2} C(u,v)\,du\,dv - \frac{1}{4})/(1/12)$
- Tau generalizes to multivariate:
  - $(\iiint cC - 2^{-n})/(\frac{1}{2} - 2^{-n})$.
  - $(\iiint C - 2^{-n})/(1/(n+1) - 2^{-n})$ for rho.

Rho and Tau Are Related (for continuous copulas, from Nelsen)



Example C(u,v) Functions

Frank: $-a^{-1}\ln(1 + g_u g_v / g_1)$, with $g_z = e^{-az} - 1$

$\tau(a) = 1 - 4/a + 4/a^2 \int_0^a t/(e^t - 1)dt$

Gumbel: $\exp\{-((-\ln u)^a + (-\ln v)^a)^{1/a}\}$, $a \geq 1$

$\tau(a) = 1 - 1/a$

HRT: $u + v - 1 + ((1 - u)^{-1/a} + (1 - v)^{-1/a} - 1)^{-a}$

$\tau(a) = 1/(2a + 1)$

Normal: $C(u,v) = B(p(u),p(v);a)$ i.e., bivariate normal applied to normal percentiles of u and v, correlation a

$\tau(a) = 2\arcsin(a)/\pi$

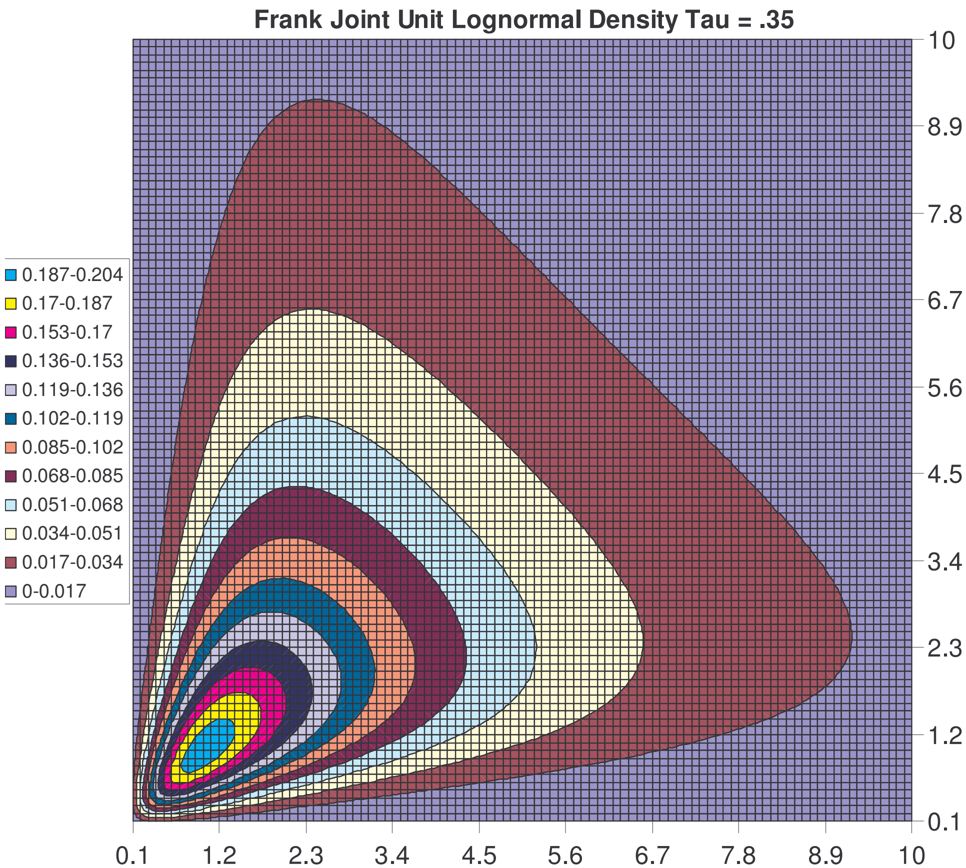
Describing Copulas

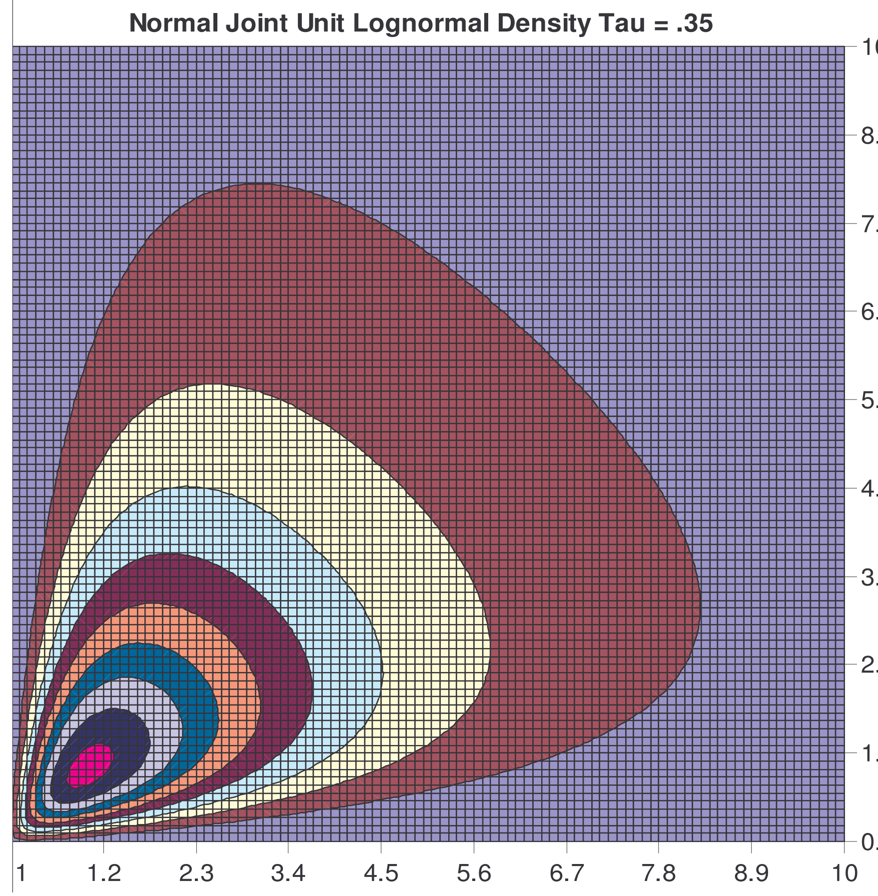
- Correlation coefficients tell you more about parameter than about copula
- Tail coefficients tell more about copula

$$R = \lim_{z \to 1} \Pr(v > z \mid u > z)$$

  - For many copulas this is zero for all parameters – e.g., normal copula
  - For others it is always positive – how positive depending on correlation
- Contours of joint distribution illuminating

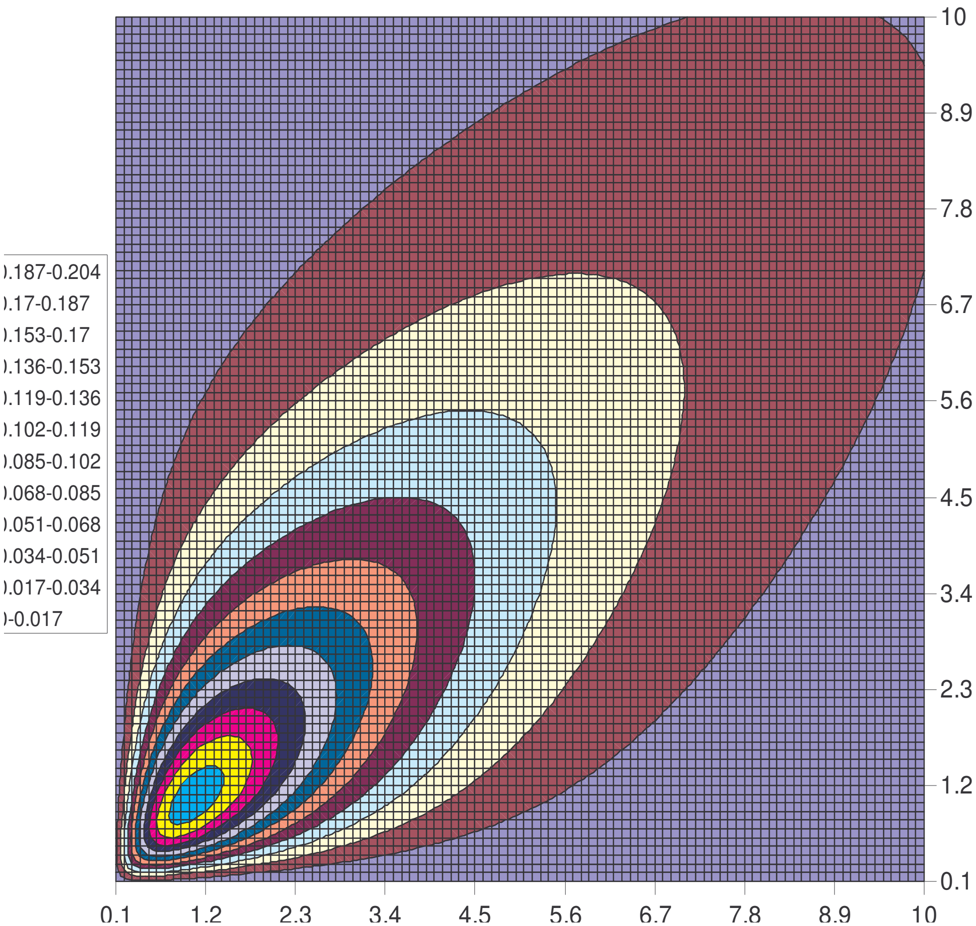
Copulas Differ in Tail Effects
Light Tailed Copulas Joint Unit Lognormal



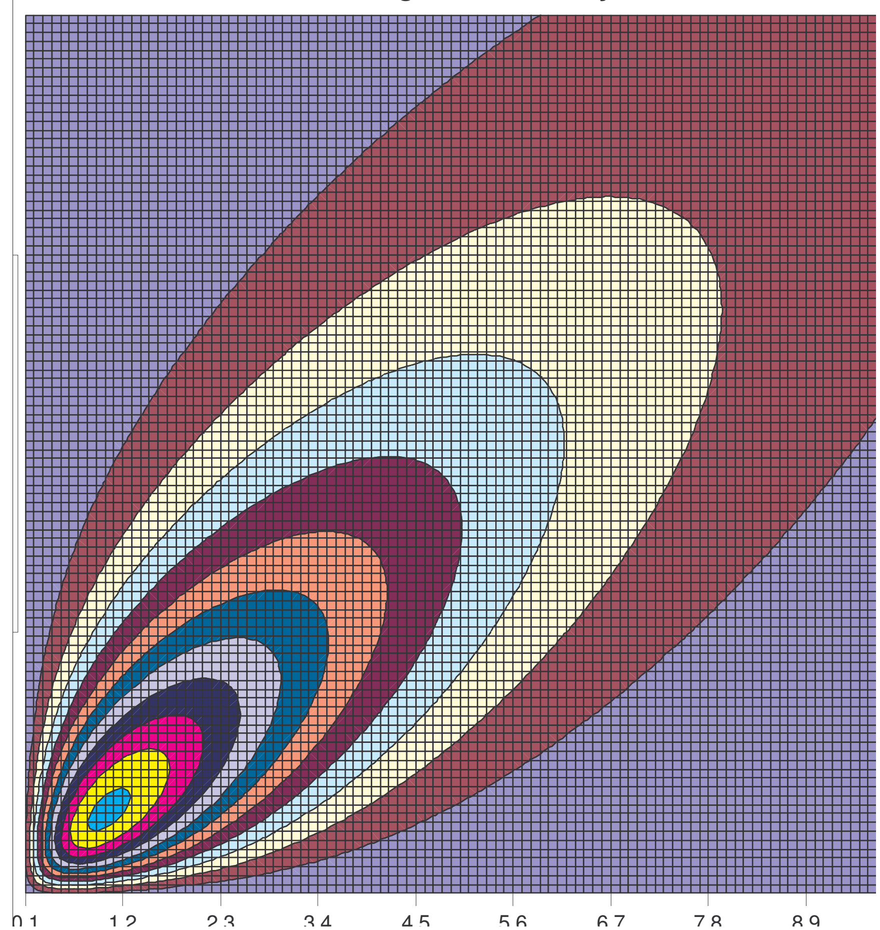


Copulas Differ in Tail Effects
Heavy Tailed Copulas Joint Unit Lognormal

Gumbel Joint Unit Lognormal Density Tau = .35



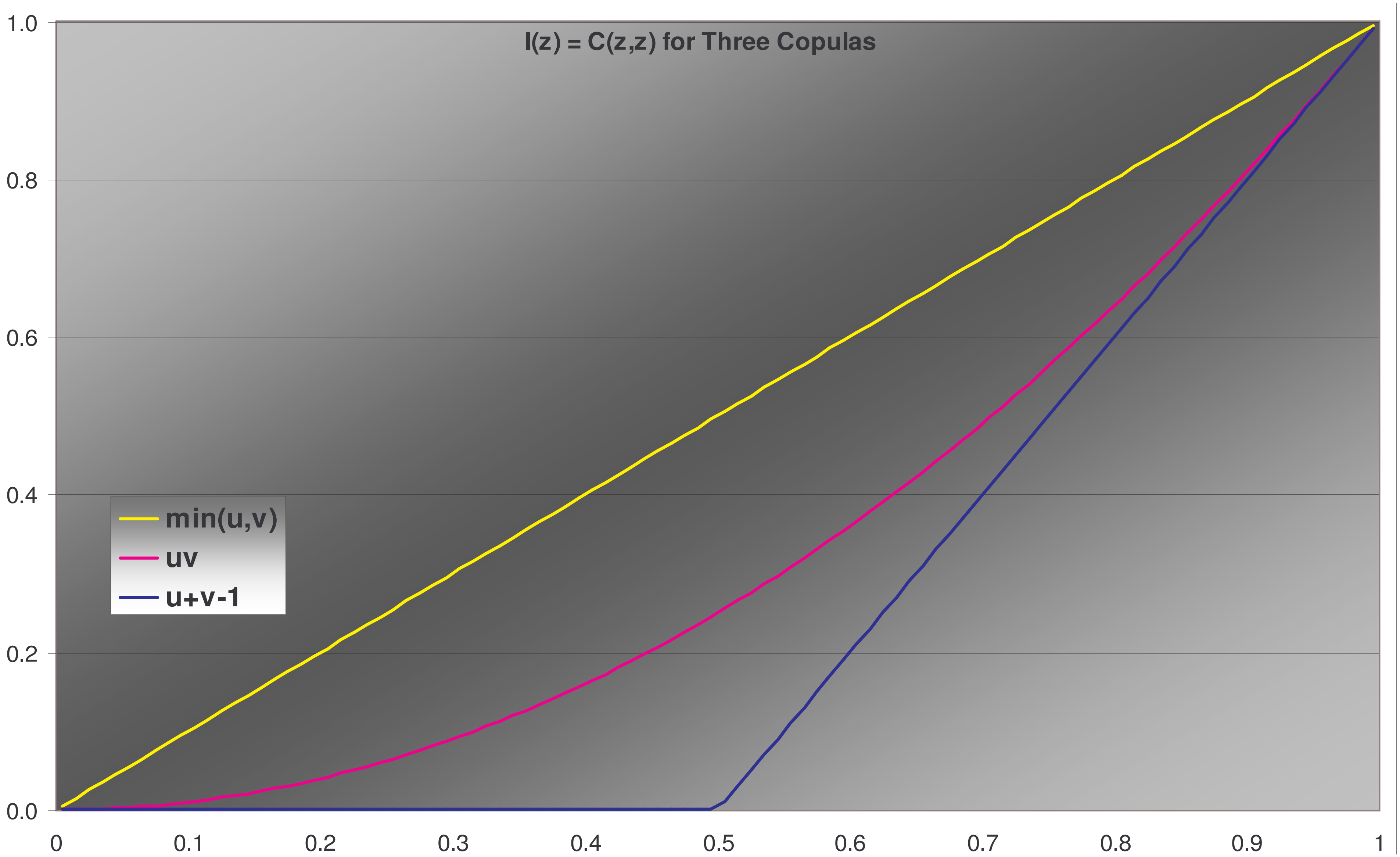
HRT Joint Unit Lognormal Density Tau = .35



Descriptive Functions

- Correlation and tail coefficients are scalar descriptors
- Descriptive functions are functions mapping $(0,1) \to (0,1)$ that describe features of the copula
- Usually $z \in (0,1)$ indicates a region of the unit square (or cube or hypercube, …)
  - E.g., $u<z$, or $u,v<z$ or $C(u,v)<z$
- Function(z) calculates a scalar on this region – often an integral over region

Copula on Diagonal I(z) = C(z,z)



Tail Concentration Functions
Venter PCAS 2002

$L(z) = Pr(U<z \mid V<z) = Pr(U<z \ \& \ V<z)/z$

$R(z) = Pr(U>z \mid V>z) = Pr(U>z \ \& \ V>z)/(1 - z)$

$L(z) = C(z,z)/z = I(z)/z$

$R(z) = [1 - 2z + I(z)]/(1 - z)$

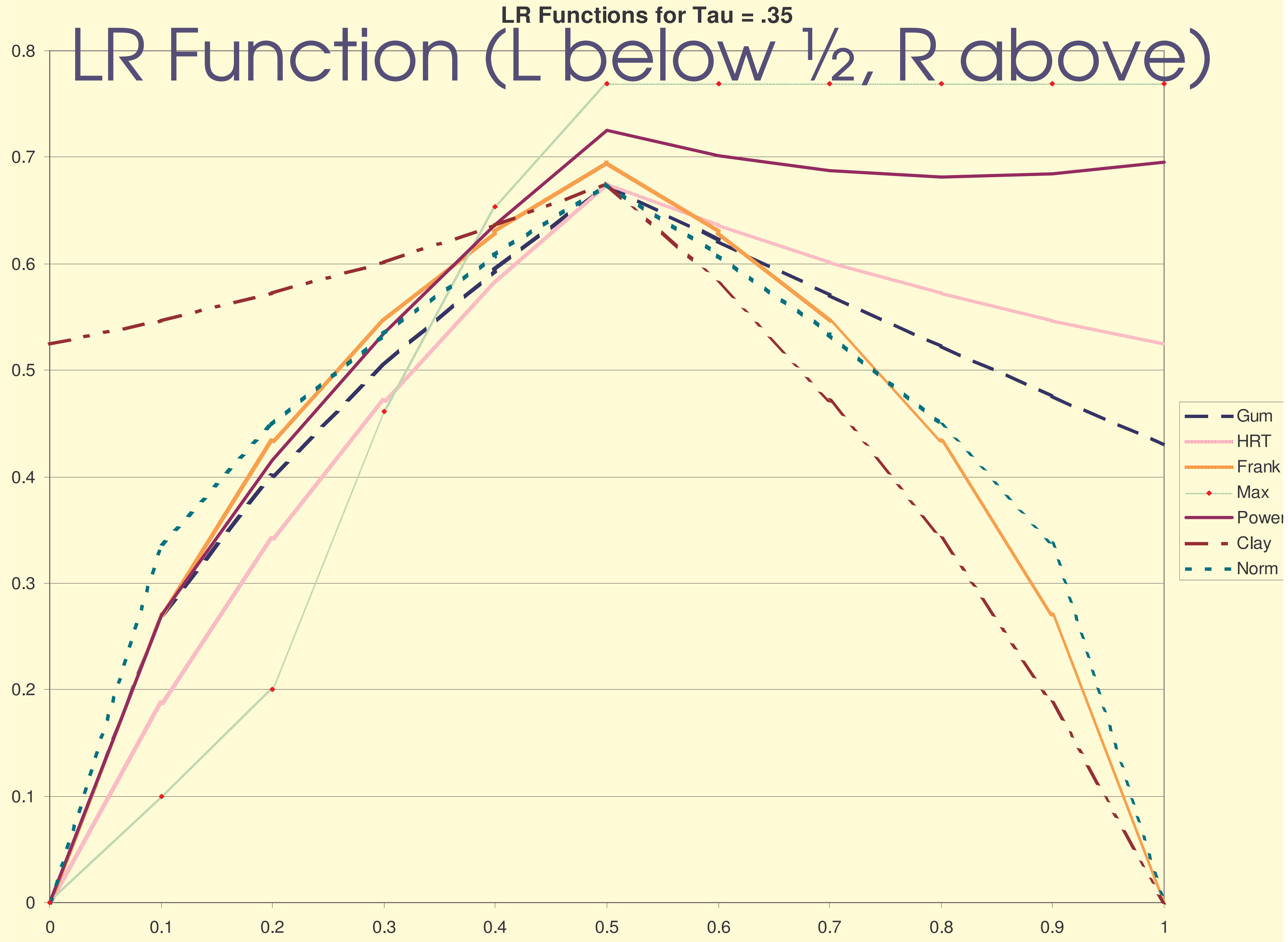
$L(1) = 1 = R(0)$

Action is in R(z) near 1 and L(z) near 0

lim R(z), z->1 is R, and lim L(z), z->0 is L

Generalizes $L(z) = Pr(U<z \ \& \ V<z \ \& \ W<z)/z$

LR Function (L below ½, R above)



Other Descriptive Functions

Tau can be calculated by

$-1+4\int_0^1\int_0^1 C(u,v)c(u,v)dvdu.$

Cumulative tau using u,v<z:

$J^-(z) = -z^2+4\int_0^z\int_0^z C(u,v)c(u,v)dvdu/C(z,z).$

$J^+(z) = -z^2+4\int_z^1\int_z^1 C(u,v)c(u,v)dvdu/C(z,z)$

$\chi(z) = 2 - \ln(C(z,z))/\ln z$

For $z \rightarrow 1$, this approaches R

Can compare empirical and fitted

Multivariate Copulas
Well Known Ones

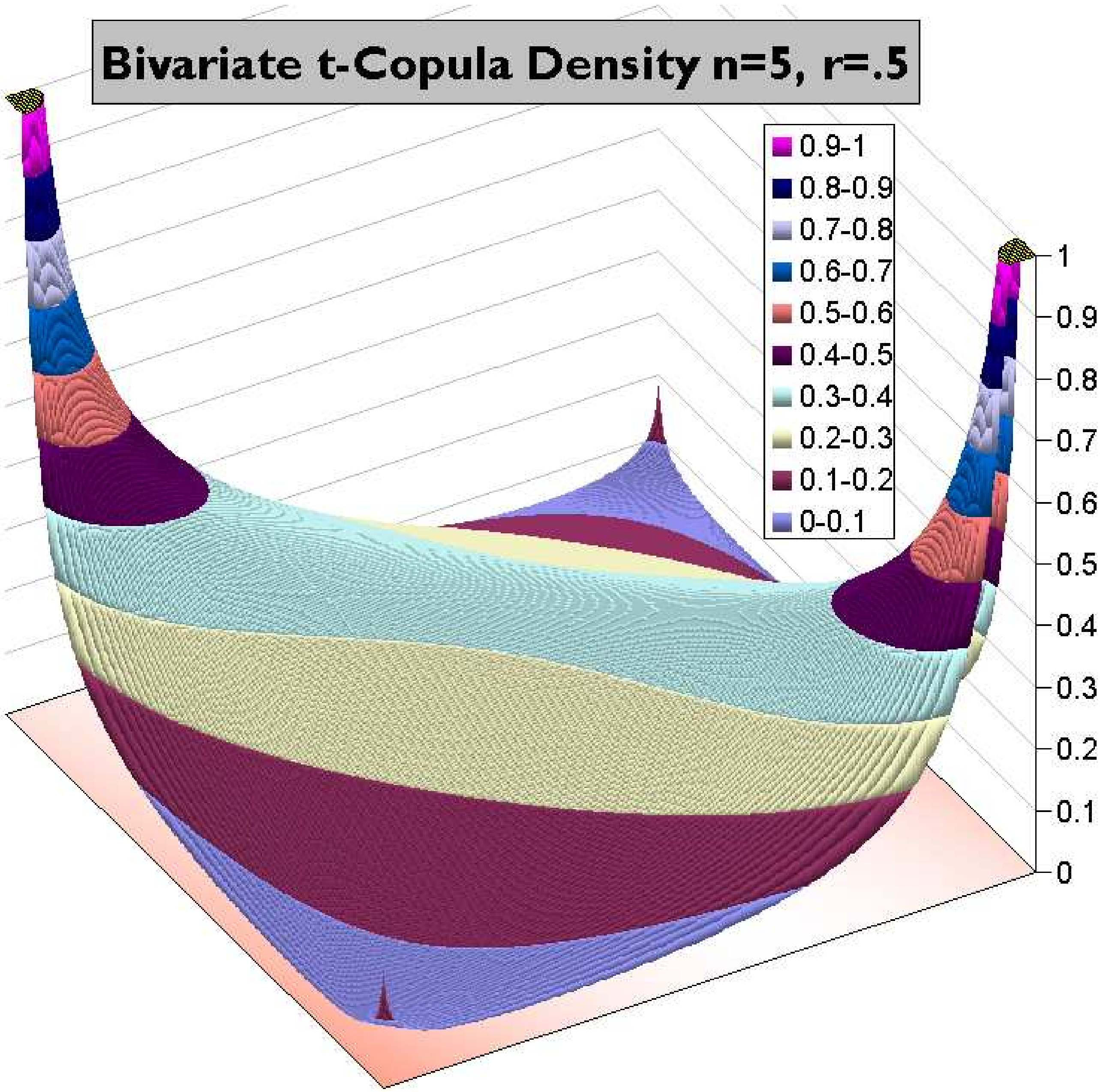
normal copula

- bivariate case: $C(u,v) = B(\Phi^{-1}(u),\Phi^{-1}(v);\rho)$

t-copula

- described best by simulation procedure:
  - simulate vector of correlated standard normals (Cholesky)
  - divide vector by function of one random draw of a gamma
  - take the t-distribution of new vector to get a vector of probabilities (gamma and t use one dof parameter)
    - as dof $\to \infty$, t goes to normal
- all hit by the same gamma draw increases tail dependence but not correlation overall

in multivariate case, correlation matrix is a set of parameters, plus one more for dof in t case


Bivariate t-Copula Density n=5, r=.5
0.9-1
0.8-0.9
0.7-0.8
0.6-0.7
0.5-0.6
0.4-0.5
0.3-0.4
0.2-0.3
0.1-0.2
0-0.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

Limitations

- Tail dependences R and L of normal = 0
- For t, $R = 2 - 2F_{n+1}\{ ((n+1)(1-\rho)/(1+\rho))^{\frac{1}{2}} \}$
  - zero (almost) for large n
  - for fixed n, increases with ρ
- So for t-copula, higher correlation has higher tail dependence
  - Reasonable but not always holding in data
- Symmetric tails

Individuated t-Copula

- A separate dof parameter for each variate
  - simulate by taking the same gamma probability draw for each variate
- Properties
  - Correlations (Kendall's etc.) close to t's
  - Two variates with same dof have same tail dependence as t

Formulas Kind of Messy

Density function:

$$c(\vec{u}) = \int_0^1 \frac{\prod_{n=1}^{N} \left\{ \sqrt{h_n(y)}\,\Gamma(\nu_n/2)\left(1+ t_n^2 / \nu_n\right)^{\frac{1+\nu_n}{2}} \Big/ \Gamma\left(\frac{1+\nu_n}{2}\right) \right\}}{\sqrt{\det(\rho)(2\pi)^N}\, \exp\left(\sum_{n,m} \frac{J_{n,m} t_n t_m}{2} \sqrt{\frac{h_n(y) h_m(y)}{\nu_n \nu_m}} \right)} dy$$

Tail dependence:

$$\int_0^\infty S(c_n y^{1/\nu_n}, c_m y^{1/\nu_m}) dy \quad c_j = \sqrt{2}\left[\Gamma\left(\frac{1+\nu_j}{2}\right) \Big/ \sqrt{4\pi} \right]^{\frac{1}{\nu_j}}$$

Reduces to t R if $\nu_n = \nu_m$. Between t values for different ν.

Other Multivariate Copulas

- MMc family defined by Harry Joe
  - Closed form copulas MM1, MM2, MM3
  - Other MMc's defined numerically
- Takes a matrix of parameters, but they are not correlations
  - One more parameter for each variable and one more overall
- R > 0, but tails not symmetric

Example: MM1

For $\delta_{ij} \geq 1$ and $\theta \geq 1$, the copula at the m-vector u is:

$$C(\mathrm{u}) = \exp\left\{ -\left[\sum_{j=1}^{m} (1-(m-1)p_j) y_j + \sum_{i<j} \left((p_i y_i)^{\delta_{ij}} + (p_j y_j)^{\delta_{ij}} \right)^{1/\delta_{ij}} \right]^{1/\theta} \right\}$$

The bivariate i, j margin is:

$$C(u_i, u_j) = \exp\left\{ -\left[(1-p_i) y_i + (1-p_j) y_j + \left((p_i y_i)^{\delta_{ij}} + (p_j y_j)^{\delta_{ij}} \right)^{1/\delta_{ij}} \right]^{1/\theta} \right\}$$

Lower tail dependence is zero. Upper tail dependence is given by:

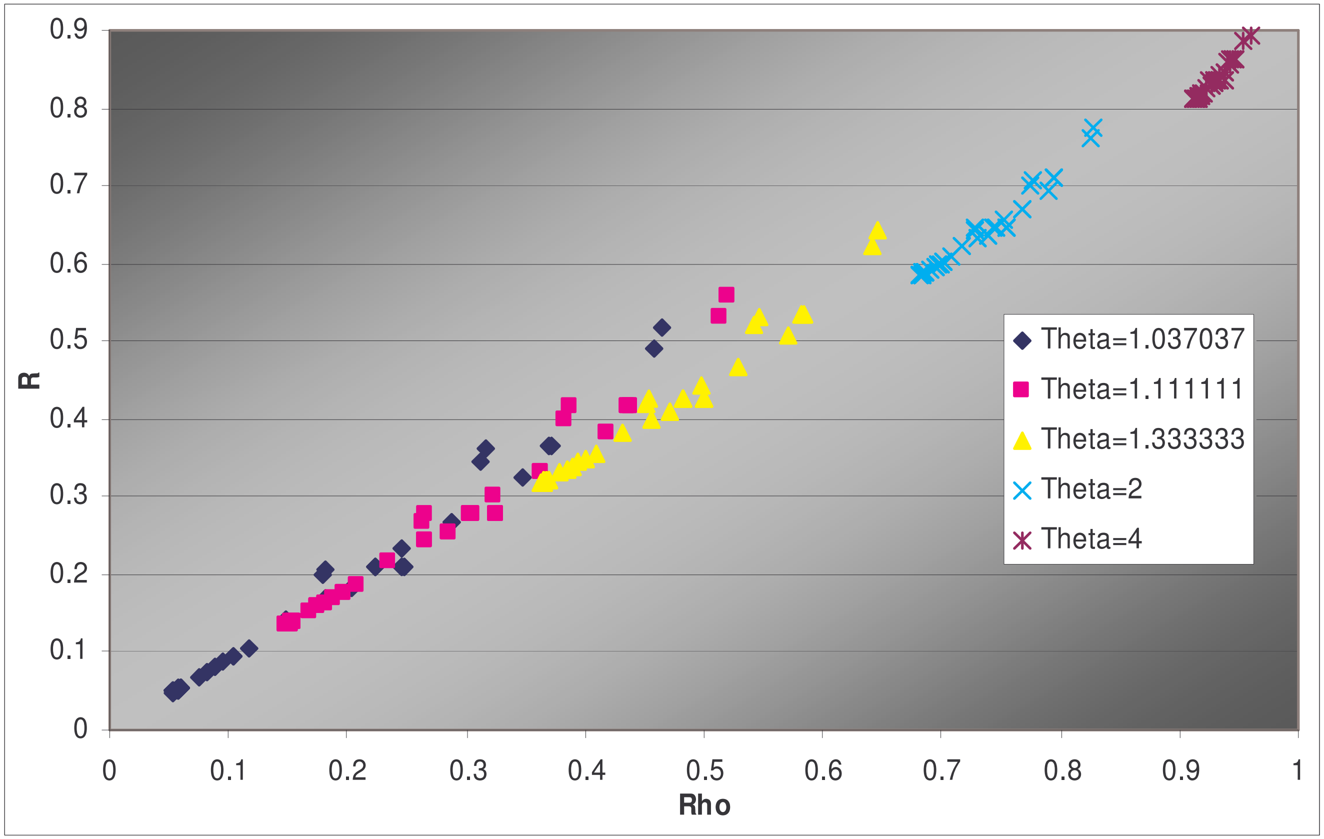
$$\lambda_{ij} = 2 - \left[2 + \left(p_i^{\delta_{ij}} + p_j^{\delta_{ij}} \right)^{1/\delta_{ij}} - p_i - p_j \right]^{1/\theta}$$

Depends on both p parameters as well as δ_{ij}

Properties of MMc Copulas

- More parameters to control correlation and tail dependence
  - Higher correlation not always with stronger tail
- Cannot match every correlation matrix
  - Either correlations should be all fairly low or all close to each other
    - More so the more variates there are
- Might work for insurance losses
  - Generally low correlations but could have high tail dependence

MM1 R as a Function of Rho by Theta, m=3



1. Once theta set, only some rho and R values possible
2. For fixed theta and rho, different R's are possible

Fitting MMc

- MLE a natural, but densities get messy
  - Especially as number of variates increases
  - Could try numerically differentiating copula to get likelihood function
- We also tried product of bivariate likelihood functions
  - Compared with MLE and not too bad

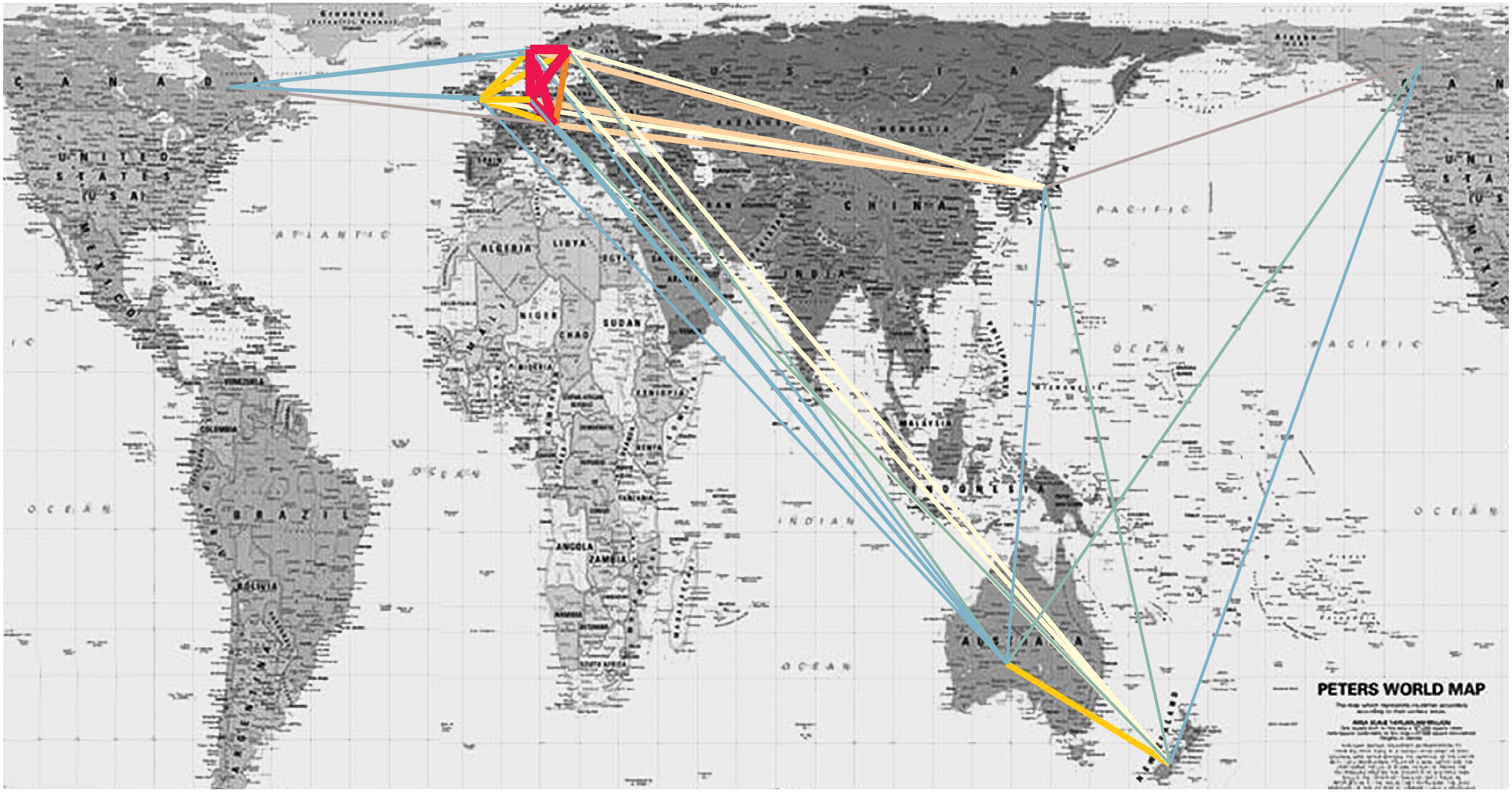
Correlation Example

Correlations for Monthly Change in Exchange Rate Against US $

1-1-71 to 1-1-007

	UK pound	Aus $	NZ $	Swiss Fr	Canada $	Japan Yen	S Kroner	D Kroner	N Kroner
UK pound	100%	25%	42%	66%	21%	44%	64%	68%	69%
Aus $	25%	100%	63%	25%	38%	29%	30%	27%	29%
NZ $	42%	63%	100%	39%	28%	38%	41%	44%	44%
Swiss Fr	66%	25%	39%	100%	16%	59%	72%	89%	80%
Canada $	21%	38%	28%	16%	100%	11%	25%	22%	24%
Japan Yen	44%	29%	38%	59%	11%	100%	45%	56%	51%
S Kroner	64%	30%	41%	72%	25%	45%	100%	81%	82%
D Kroner	68%	27%	44%	89%	22%	56%	81%	100%	88%
N Kroner	69%	29%	44%	80%	24%	51%	82%	88%	100%

10%	19%
20%	29%
30%	39%
40%	49%
50%	59%
60%	69%
70%	79%
80%	89%



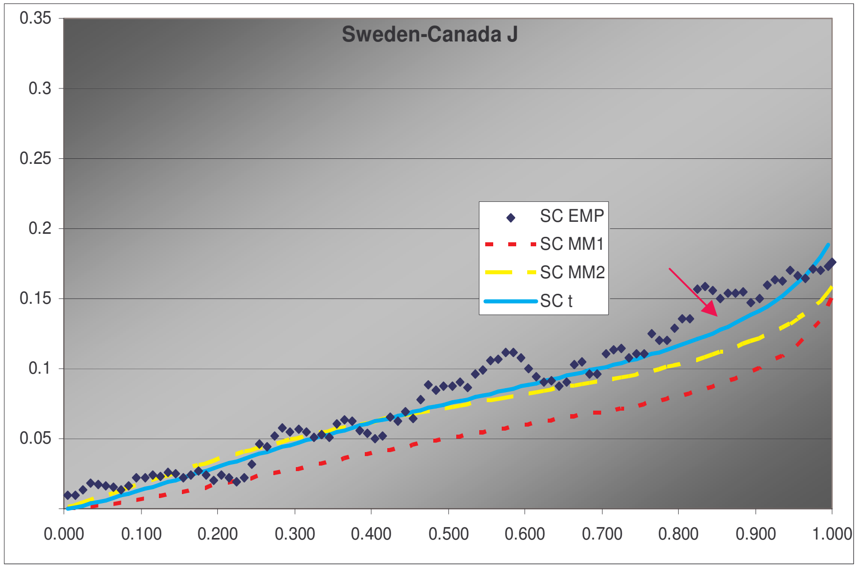
Fitting to Canada, Japan, Sweden through September '005

	MM2			MM1				t	
	SSE	MLE bi		SSE	MLE bi	MLE tri		MLE bi	MLE tri
δ_{12}	2.62588	1.50608		3.69513	2.14275	2.1094		0.491	0.490
δ_{13}	0.80055	0.43963		1.52005	1.19832	1.1152		0.262	0.266
δ_{23}	3.2E-07	0.0103		1	1	1		0.097	0.097
p_1	0.49881	0.37649		0.49963	0.40533	0.37175			
p_2	0.5	0.5		0.5	0.5	0.5			
p_3	0.26236	0.49625		0.29097	0.5	0.5			
θ	0.19599	0.2209		1.08308	1.0939	1.1234		20.53	20.95

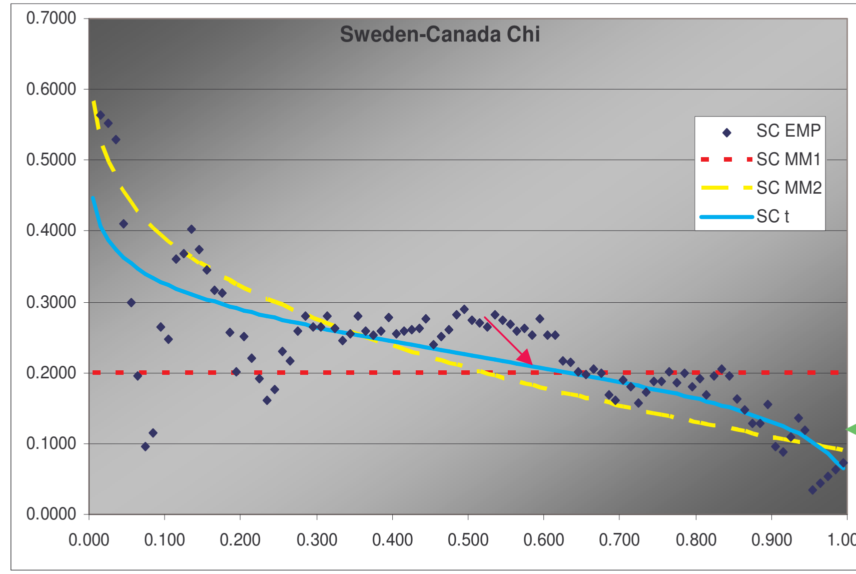
MLE trivariate and product of bivariate likelihood functions parameters similar (MM1 & t)

SSE parameters quite different from MLE (MM1 & MM2)

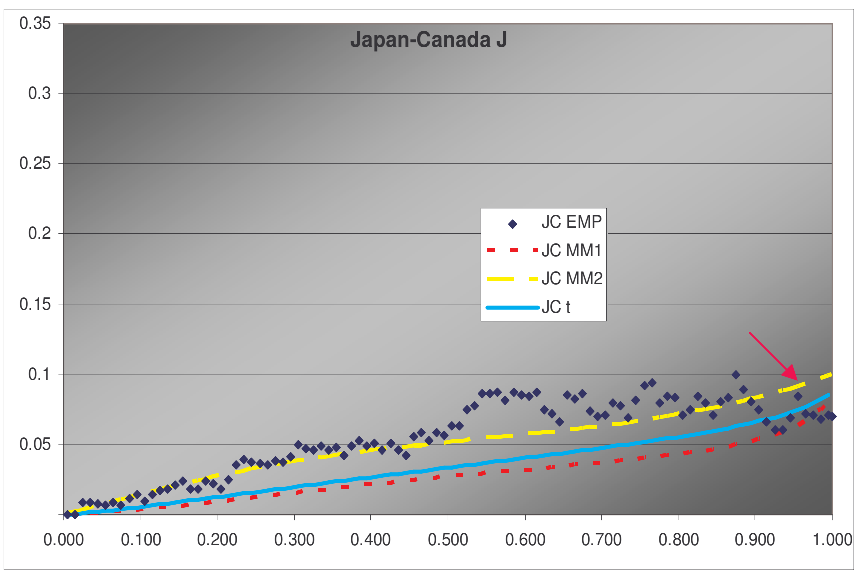
Range of correlations a stretch for MMc



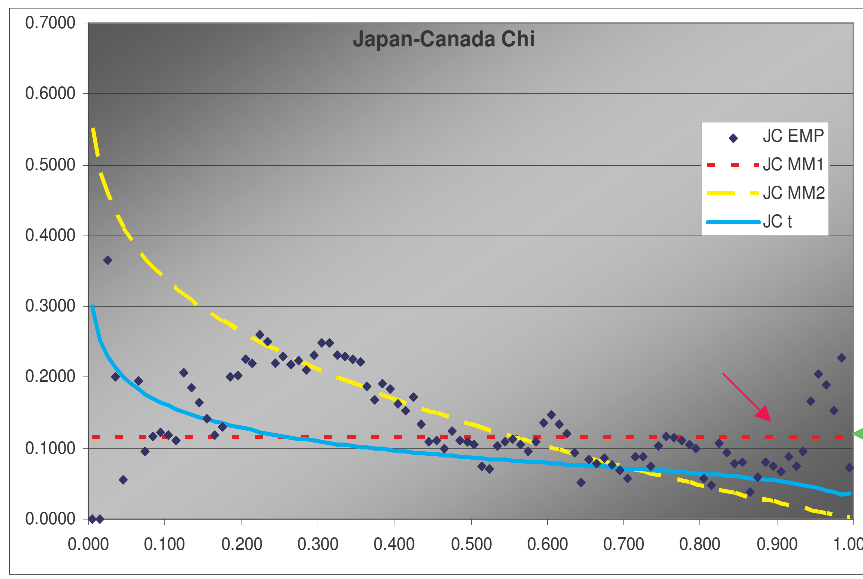
Sweden – Canada J Function



Sweden – Canada χ Function



Japan – Canada J Function



Japan – Canada χ Function

Summary

- MMc's and t have different shapes
- IT and MMc give more control over tail than t
- MMc's work better when correlations small or close to each other, possibly as in insurance lines