The Known, the Unknown, and the Unknowable: Challenges in Validating AMA Models

Emre Balta, Office of the Comptroller of the Currency

The views expressed in this presentation are solely those of the author and do not necessarily reflect official positions of the Comptroller of the Currency and the Department of the Treasury.

AMA – Regulatory Requirements

- Basel II does not stipulate a specific approach or distributional assumptions.
- However, an AMA model should meet the following minimum regulatory requirements:
  - quantify risk at the 99.9th confidence level over a one year horizon
  - minimum five years of data
  - capture potentially severe tail events
  - granularity of the model – commensurate with the risk profile of the bank
  - incorporate dependence
  - appropriately weight each of the 4 "elements": internal data, external data, scenarios and BEICFs

Nature of Operational Loss Distributions

1. Operational loss distributions are:
   1. extremely heavy-tailed
   2. extremely skewed
   3. have special dependence characteristics
2. Operational loss data exhibits heterogeneity.

Extreme caution must be taken when modeling heavy-tailed processes.

Nature of Operational Loss Data

Table ILD6
Distribution of Loss Amount by Severity of Loss

Severity of Loss	Cross-Bank Median of Distribution Across Severity Brackets			Total	
	Number of Losses	Gross Loss Amount	Gross Loss Amount Net of Non-Insurance Recoveries	Number of Losses	Gross Loss Amount (€Millions)
€0 ≤ X < €20,000	91.29 %	26.26 %	18.86 %	9,897,083	12,164
€20,000 ≤ X < €100,000	6.52 %	12.63 %	15.66 %	121,533	5,178
€100,000 ≤ X < €1 Million	1.83 %	19.37 %	21.35 %	30,598	8,085
€1 Million ≤ X < €2 Million	0.15 %	5.48 %	6.12 %	1,688	2,401
€2 Million ≤ X < €5 Million	0.12 %	9.05 %	9.10 %	1,116	3,570
€5 Million ≤ X < €10 Million	0.04 %	6.87 %	7.90 %	404	2,827
€10 Million ≤ X < €100 Million	0.04 %	15.55 %	17.39 %	333	8,243
€100 Million ≤ X	0.02 %	41.79 %	43.51 %	41	21,752
All				10,052,796	64,221

Note 1. X = severity of loss, based on gross loss net of non-insurance recoveries.
Note 2. All losses in the stable dataset.
Note 3. Results for losses less than €20,000 are not complete as loss data collection thresholds differ across participants.
Note 4. Median calculations include only banks with losses in each particular severity category. If a bank reports no losses in a category, it is not included in the calculation.

Source: BCBS (2009)

Model Risk

- A model, by definition, is an abstraction of the real world.
- All models are conditional on data.
- A model may be expressed in two parts as (Draper, 1995).

$$M = (S, \theta)$$

  where

  - S : sets of structural assumptions, and
  - θ : parameters whose meaning is specific to the chosen structure(s).

- Models are generally used in two forms:
  1. As representaion of what's going on
  2. As description of the data

Model Risk

Three main sources of uncertainty in the model building process (Chatfield, 1995):

1. model uncertainty: uncertainty about the structure of the model;
2. parameter uncertainty: uncertainty about estimates of the model parameters, assuming that we know the structure of the model;
3. unexplained random variation in observed variables even when we know the structure of the model and the values of the model parameters.

Model Risk

"It is hard for us, without being flippant, to even see a scenario within any kind of realm of reason that would see us losing a dollar in any of those transactions."

- Joseph Cassano, head of AIGFP, at a conference call to AIG investors

(Bloomberg Markets, August 2009)

Model Risk

In the context of banking and supervision, model risk describes the potential for losses that might be sustained due to

- the use of misspecified models (e.g., mis-specified tail behavior)
- the use of erroneously calibrated models
- an overly broad interpretation or use of a model beyond the scope of application for which it was developed.

Validation Process

A complete validation process is more than statistical testing. It consists of:

- a timetable for validation activities
- identification of parties responsible for validation
- tests and analyses to be performed
- actions to be taken in response to findings
- documentation and reporting of findings

Validation Challenges

Model uncertainty

- Challenges for both bankers and regulators
  - Understanding underlying assumptions and their appropriateness
  - Understanding the model's theoretical soundness and mathematical integrity
  - Sensitivity analysis (under stress conditions)
- Challenges particularly for regulators
  - Flexibility inherent in the rule creates a broad range of practice:
    - Stochastic ordering of loss models (e.g., Which particular aspects of the model determine the first-order outcome?
    - Understanding the root cause of the dispersion
  - Developing credible benchmarks for operational risk exposure

Validation Challenges

Broad range of practice, evidently, results in over-dispersion of results

Table C1
Reported Regulatory Operational Risk Capital as a Percentage of:
Assets, Tier 1 Capital and Gross Income

			All	Australia	Europe	Japan	North America	Brazil / India
Consolidated Assets	All	Median (25th-75th)	0.33% (0.24%-0.47%)	0.30% (0.27%-0.37%)	0.34% (0.18%-0.46%)	0.25% (0.22%-0.28%)	0.38% (0.33%-0.58%)	0.46% (0.44%-0.54%)
	AMA	Median (25th-75th)	0.27% (0.20%-0.40%)	0.29% (0.27%-0.30%)	0.23% (0.17%-0.37%)	0.20% (0.20%-0.23%)	0.49% (0.33%-0.80%)	na na
	Non-AMA	Median (25th-75th)	0.38% (0.29%-0.49%)	0.35% (0.26%-0.41%)	0.38% (0.27%-0.47%)	0.28% (0.24%-0.29%)	0.35% (0.33%-0.38%)	0.46% (0.44%-0.54%)
Consolidated Tier 1 Capital	All	Median (25th-75th)	7.51% (5.21%-10.25%)	7.11% (6.39%-7.77%)	7.97% (5.48%-10.21%)	4.88% (4.19%-5.96%)	9.18% (6.07%-11.51%)	8.02% (4.42%-9.64%)
	AMA	Median (25th-75th)	7.38% (5.30%-9.63%)	7.25% (7.11%-7.70%)	8.91% (6.53%-9.78%)	4.46% (3.51%-5.44%)	10.90% (5.45%-12.40%)	na na
	Non-AMA	Median (25th-75th)	7.62% (5.21%-10.58%)	6.72% (5.03%-9.84%)	7.66% (5.23%-10.46%)	5.13% (4.73%-6.06%)	8.53% (7.62%-9.82%)	8.02% (4.42%-9.64%)
Consolidated Gross Income	All	Median (25th-75th)	12.27% (10.58%-14.96%)	10.06% (4.46%-14.49%)	12.09% (10.72%-13.63%)	14.05% (13.08%-14.86%)	12.65% (8.59%-17.37%)	7.53% (5.19%-12.50%)
	AMA	Median (25th-75th)	10.83% (8.38%-13.83%)	7.82% (3.83%-10.06%)	10.70% (9.47%-13.36%)	12.44% (11.53%-13.39%)	11.63% (6.67%-21.76%)	na na
	Non-AMA	Median (25th-75th)	12.79% (11.33%-15.03%)	13.86% (6.96%-18.10%)	12.10% (11.42%-14.08%)	14.58% (14.00%-14.92%)	13.08% (10.69%-13.87%)	7.53% (5.19%-12.50%)

Note 1. 25th-75th represents the interquartile range, which is the range of values (between the 25th percentile and 75th percentile) that contains half the banks in the sample.

Note 2: All participants in Brazil / India are non-AMA.

Problem Formulation

- Notation
  - $X_k^i -$ k^{th} loss for a unit of measure (UOM) i
  - N^i $-$ frequency process
  - L_i $-$ aggregate loss for UOM i
- Given the frequency process N^i and the severity process X_k, the aggregate loss L_i for the unit of measure i is a compound process defined as:
$$L_i = \sum_{k=1}^{N^i} X_{i,k} \tag{1}$$
where $X_{i,k}$ are assumed to be *iid* within a UOM, and independent of N.
- (Firmwide) aggregate loss for a bank with d units of measure is then:
$$L_D = \sum_{i=1}^{d} L_i = L_1 + \ldots + L_D \tag{2}$$

Problem Formulation - Aggregate Loss Distribution

- Let G be the *df* of the random sum L_D and X be *iid* positive random variables. Then, the aggregate loss distribution is denoted by

$$G_{L_D}(x) = P\left(L_D \leqslant x\right) \tag{3}$$

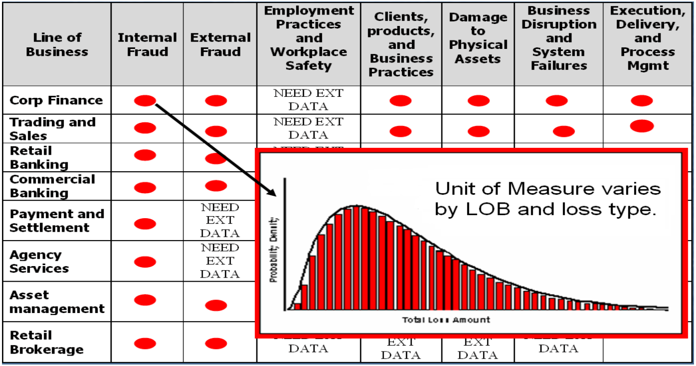
  - G_{L_D} has complicated/unknown structure
  - If we had a multivariate model in the form of a joint *df*, then the dependence structure would be implicitly described.

- The minimum regulatory capital (MRC) is calculated as a quantile of the G_{L_D} at the confidence level α=0.999:

$$VaR_{0.999}\left(_{L_D}\right) = G^{\leftarrow}_{L_D}(0.999) = \inf\left\{x \in \mathbb{R} \mid G_{L_D}(x) \geqslant \alpha\right\} \tag{4}$$

Problem Formulation - Schematic Representation

- Loss distributions vary across loss types and business lines.
- Assuming that a bank uses Basel II Level 1 event type and business line classification in its UOM scheme, it will have, at most, a 56-dimensional problem.



Modeling Alternatives

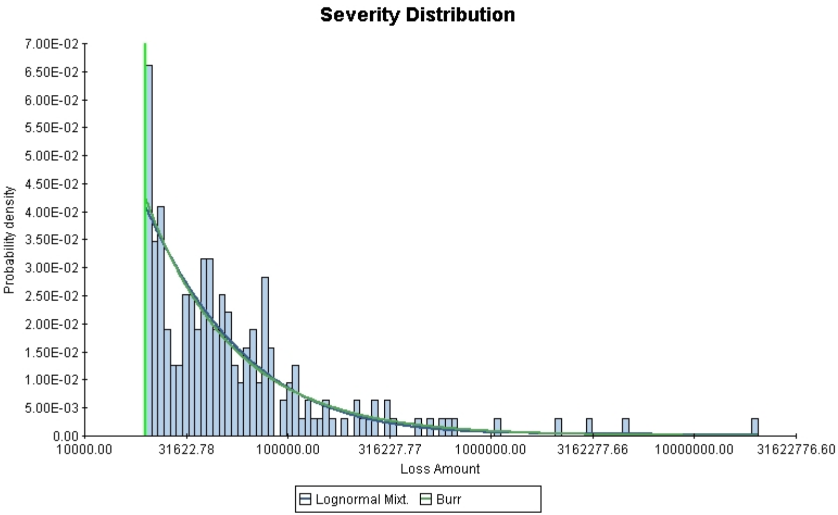
Op risk models could broadly be classified as different variants of the

- Loss Distribution Approach (LDA). LDA-based models are very well known from the (non–life) actuarial literature; see Klugman et al. (2004), Frachot et al. (2001, 2004)
- Extreme Value Theory (EVT). Embrechts et al. (1997), McNeil et al. (2005).
- Approximate analytical formulas. Böcker and Klüppelberg (2005), Böcker (2006) and Böcker & Spritulla (2006).

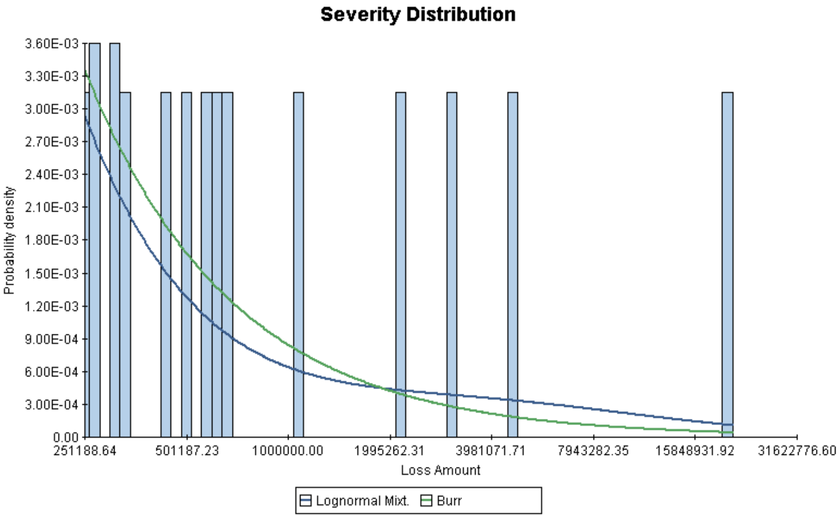
AMA in Practice - Building Blocks

1. Pool the data by business line or event type or a combination of both
2. Fit frequency and severity distributions
3. Get aggregate distribution through convolution
4. Estimate $\widehat{VaR_1}, \ldots, \widehat{VaR_n}$
5. Add (comonotonicity): $\widehat{VaR_{firm}} = \sum_{i=1}^{n} \widehat{VaR_i}$
6. Take diversification benefit: $VaR_{reported} = (1 - \delta)\,\widehat{VaR_{firm}}$

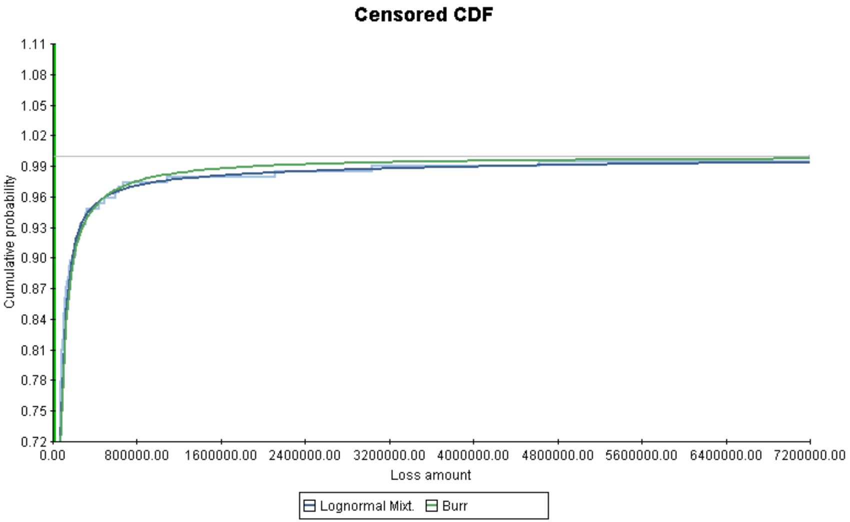
Example: Severity fit - Mixture lognormal vs. Burr distribution



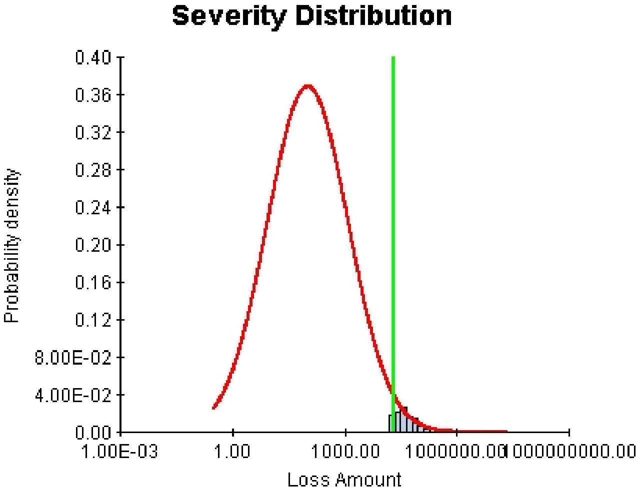
Example: Severity fit - Mixture lognormal vs. Burr distribution (tail)



Example: Severity fit - Mixture lognormal vs. Burr distribution



Example: Impact of data truncation on the severity fit



Example: Impact of severity choice on capital

- Both the mixture lognormal and the Burr distributions fit body of the data well
- However, each result in quite different tail estimates:

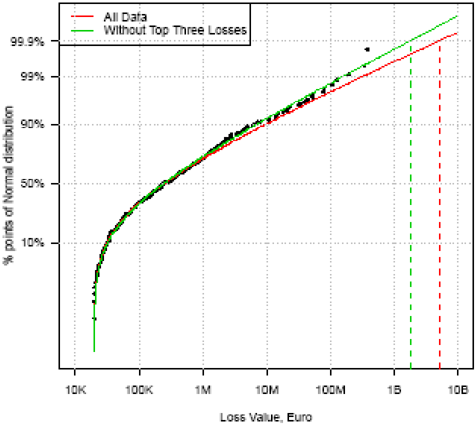
	Mixture lognormal	Burr
VaR(0.999)	122,963,094	426,208,494

Volatility of Capital Estimates

- Extreme losses dominate: An extreme loss can cause drastic changes in the estimate of capital.
- Rare event prediction: Statistical estimation of this uncertainty (95% confidence levels) are typically very wide for the estimation of rare events.
- Need significantly more data to reduce standard errors

Volatility of Capital Estimates

Example: (Cope et al., 2009) Lognormal fits to ORX Corporate Finance / Clients, Products, and Business Practices loss data, using all data vs. excluding the three largest losses



Dominance of Sums

- Value of the most extreme losses determine the aggregate annual loss distribution.
- For heavy-tailed dist, the following theoretical result holds for the high quantiles of the aggregate loss distributions: tail of the maximum determines the tail of the sum.

$$\lim_{x\to\infty} \frac{P(X_1 + ... + X_n > x)}{P\left(\max\left(X_1, ..., X_n\right) > x\right)} = 1 \qquad \textit{and} \qquad \lim_{x\to\infty} = \frac{\overline{F^{n*}}(x)}{\overline{F}(x)} = n$$

- The operational risk capital is driven by a few low-probability, high-severity events, rather than the accumulation of many *iid* high-probability, low-severity events (ie, F is a subexponential distribution function)

Dominance of Sums - Analytical Approximation

For subexponential distributions, Bocker and Kluppelberg (2005) derive a closed-form approximation for the operational VaR, valid at very high confidence levels

Theorem

Aggregate loss distribution of *iid* **subexponential rvs:** *Let* F^{n*} *be the df of the of the random sum* $\sum_{i=1}^{n} x_i$ *and* X_i *be iid positive random variables with df* F *such that* $F(x) < 1$ *for all* $x > 0$. *Then*

$$\overline{F^{n*}}(x) \sim n\overline{F}(x), \qquad x \to \infty$$

where $\overline{F}(.) = 1 - F(.)$ *and* $\overline{F^{n*}}(.) = 1 - F^{n*}$ *are the tail distributions of severity and aggregate loss, respectively.*

Analytical Approximation

Closed-form approximation for operational VaR: Given the aggregate loss distribution G_t, the VaR for the time interval t at the confidence level κ is defined as the κ-quantile of the aggregate loss distribution:

$$VaR_t(\kappa) = G_t^{\leftarrow}(\kappa) = F^{\leftarrow}\left(1 - \frac{1-\kappa}{EN(t)}\right), \kappa \in (0,1)$$

Pros

- VaR at high level of α is independent of the body of the distribution
- It is independent of all characteristics of the frequency distribution other than its expectation

Cons

- Practical value of approximation depends on the magnitude of the approximation error.
- It underestimates the quantile by an amount which grows with the mean frequency $EN(t)$ (Mignola & Ugoccioni, 2006)

- Note from Eq. (4) that $MRC = VaR_{0.999}(_{L_D}) = VaR_{0.999}(\sum_{i=1}^{d} L_i)$
- However, as we noted before, G_{L_D} has, often, an intractable structure.
- To get around this issue, in practice, banks aggregate quantiles (VaRs) across units of measure and allow for a diversification benefit $\delta\ (\geq 0)$:

$$MRC = VaR_{0.999}\left(\sum_{i=1}^{d} L_i\right) = (1-\delta)\sum_{i=1}^{d} VaR_{0.999}\left(L_i\right)$$

- However, unless the dependence assumptions underlying $\delta(>0)$ are sound, and are robust to a variety of scenarios, implemented with integrity, and allow for the uncertainty surrounding the estimates, Basel II requires the diversification benefit δ to be zero.

Dependence - Multivariate Extremal Behavior

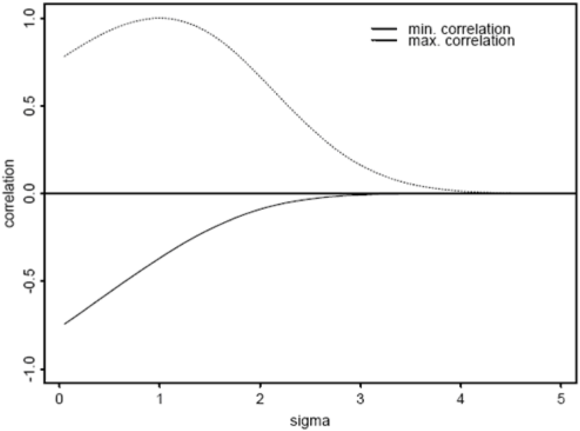
- Outside the realm of the *"normal"* world, multivariate behavior, especially in the tail region, needs to be better understood.
- For the class of elliptical models (eg, normal distribution), questions concerning diversification & capital allocation are well understood:
    - In the elliptical world, VaR is sub-additive, meaning that the VaR of a sum of risks is bounded above by the sum of the individual VaRs
    - In the non-elliptical world, VaR is no longer a coherent risk measure (Artzner *et al.*, 1999); the sum of individual VaRs does not constitute an upper bound.
- The diversification benefit δ is not constant but depends on the quantile level.

Dependence - Shortcomings of Linear Correlation

- The correlation $\rho(X_1,X_2)$ is a measure of linear dependence; $\rho \in [-1, 1]$.
- If X_1 and X_2 are independent, $\rho(X_1,X_2) = 0$, but the converse is not true.
- Correlation is only defined when the variances of X_1 and X_2 are finite.
- McNeil et al. (2005, Theorem 5.25) show that "attainable" correlations can form a strict subset of the interval $[-1, 1]$, $[\rho_{\min}, \rho_{\max}]$ with $\rho_{\min} < 0 < \rho_{\max}$:
  - The min correlation $\rho = \rho_{\min}$ is attained iff X_1 and X_2 are countermonotonic. The max correlation $\rho = \rho_{\max}$ is attained iff X_1 and X_2 are comonotonic.
  - $\rho_{\min} = -1$ iff X_1 and $-X_2$ are of the same type, and $\rho_{\max} = 1$ iff X_1 and X_2 are of the same type.

Dependence - Max and Min Corr with Marginal Lognormals

Max and min attainable correlations for lognormal rvs X_1 and X_2, where $\log(X_1)$ is standard normal and $\log(X_2)$ has mean zero and variance σ^2 :



The following conclusions can be drawn from the previous graph (where $\sigma_{x_1} = 1$ and σ_{x_2} varies across σ):

- Both positive and negative linear correlation measures vary with σ of the second distribution; this emphasizes the fact that linear correlation depends on the parameter σ of the marginal distribution.
- It is possible to attain perfect positive correlation ($\rho = 1$) iff $\sigma_{x_2} = 1$.
- For higher σ_{x_2}, the $\rho_{\max}$ moves progressively lower than 1 as the σ_{x_2} rises.
- A positive linear correlation is impossible once σ_{x_2} exceeds approximately 4. This does NOT imply, however, that the two distributions are independent despite a linear correlation of 0.

Copulas and Tail Dependence

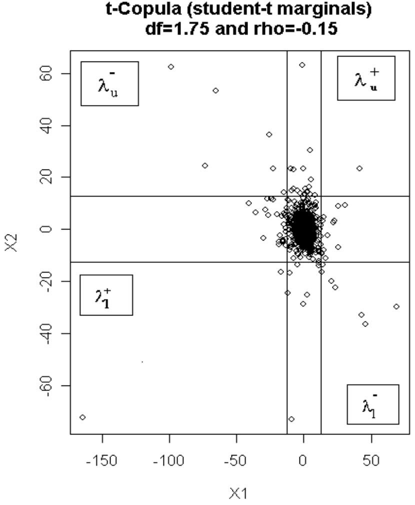
- Tail dependence measures do not depend on the underlying marginals and hence can be written in terms of the joint distribution's underlying distributional copula.
  - Positive Tail Dependency λ^+ is the probability that within the loss year both variables will exhibit similar behavior in the tails. There are two types of positive tail dependency:
    - Positive upper-tail dependence λ_u^+: the probability that both rvs will produce similar large losses in the upper tail;
    - Positive lower-tail dependence λ_l^+ : the probability that both rvs will produce similar small losses in the lower tail
  - Negative Tail Dependency λ^- is the probability that within the loss year both variables will exhibit opposite behavior in the tails. This concept is a major driver of diversification in market risk since it implies that very large losses in one tail can be offset by very large gains in the opposite tail:

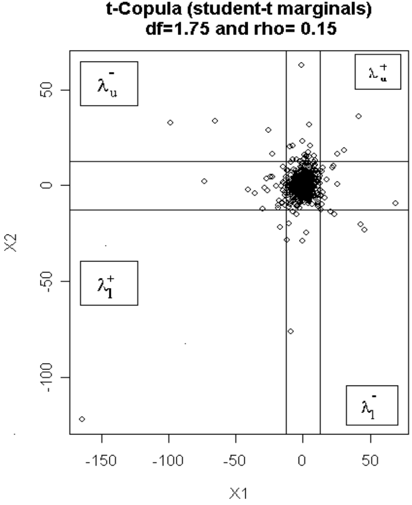
Copulas and Tail Dependence

- A copula may have
  - both upper and lower tail dependence (t-copula, Generalized Clayton copula)
  - upper tail dependence only (Gumbel copula)
  - lower tail dependence only (Clayton copula)
  - no tail dependence (Gaussian copula, Frank copula)
- Effective negative tail dependency is reliant upon two-sided data support. Operational-risk losses have only one-sided support ($X \in (o, \infty)$) since "negative" operational-risk losses are undefined.
- Major diversification benefits are provided by copulas exhibiting negative tail dependencies when the pair of X variables has two-sided support: large positive values from one variable can offset large negative values from the other variable.

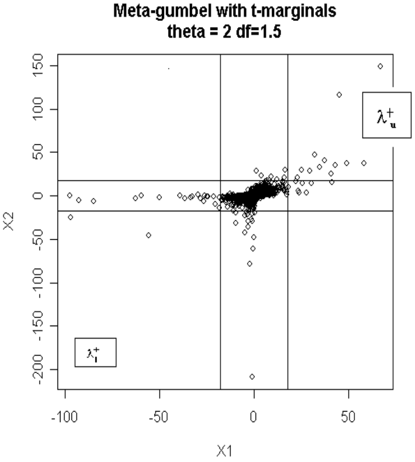
Copulas and Tail Dependence

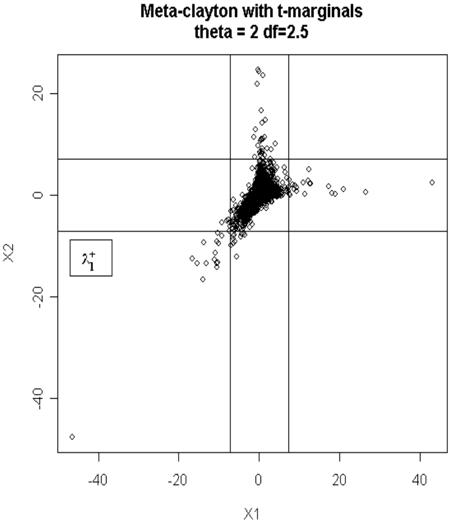
Example: (Ulman & Inanoglu, 2010) Two bivariate t-copulas with identical underlying t-marginals



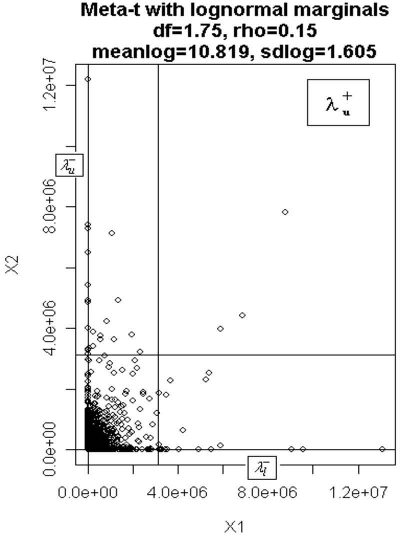


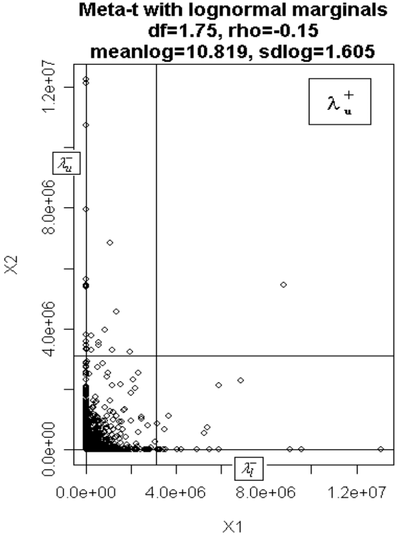
Example: (Ulman & Inanoglu, 2010) Two-sided support is not always sufficient for negative tail dependence



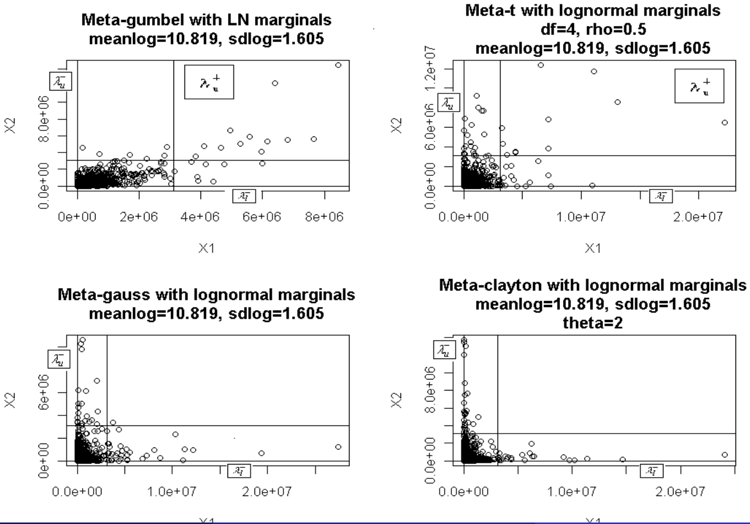


Example: (Ulman & Inanoglu, 2010) Absence of two-sided support implies significantly lower "diversification" benefits





Example: (Ulman & Inanoglu, 2010) Simulated losses from our different copulas each with identical lognormal marginals



Thank You!