Chapter 18

Dependence

Chapter 16 left a gap: single-name curves pin every marginal default distribution and say nothing at all about the joint one. This chapter is about what to put in that gap. It builds the tools that describe dependence — Sklar's theorem, rank correlation, and the tail dependence coefficient that turns out to be the number most portfolio instruments are actually a bet on — and then gives four things to use, in increasing order of how much structure they demand. The Gaussian copula appears last, as a diagnosis rather than a target: it is a coherent model whose one defining property is that it sets to zero the exact quantity a senior tranche is priced off, and we finish by measuring what that costs.

18.1 The Gap, Stated Precisely

A basket of n names, each with a default time τ_i whose distribution F_i chapter 16 bootstrapped from that name's own CDS quotes. Anything depending on more than one name — a first-to-default swap, a tranche, the counterparty exposure of chapter 24 — depends on the joint distribution $F(t_1, \ldots, t_n)$, and the marginals do not determine it.

The following theorem says exactly how much freedom that leaves.

Theorem 18.1 (Sklar)**.** *For any joint distribution F with marginals $F_1, \ldots, F_n$ there is a function $C : [0,1]^n \to [0,1]$ with*

$$F(t_1, \ldots, t_n) = C\big(F_1(t_1), \ldots, F_n(t_n)\big), \tag{18.1}$$

and C is itself a distribution function on the unit cube with uniform marginals. It is unique wherever the marginals are continuous. Conversely, any such C combined with any marginals gives a valid joint distribution.

Proof. With continuous marginals the copula is not merely shown to exist, it is written down.

The probability integral transform is the whole idea: if F_i is continuous then $U_i = F_i(\tau_i)$ is uniform on $[0,1]$, since $\mathbb{P}(F_i(\tau_i) \le u) = \mathbb{P}(\tau_i \le F_i^{-1}(u)) = u$. So define

$$C(u_1, \ldots, u_n) = F\big(F_1^{-1}(u_1), \ldots, F_n^{-1}(u_n)\big), \tag{18.2}$$

which is the joint distribution of $(U_1, \dots, U_n)$. It is a distribution function because F is; its marginals are uniform by the transform; and substituting $u_i = F_i(t_i)$ into (18.2) recovers (18.1). Uniqueness is immediate from the same substitution: (18.1) determines C at every point of the form $(F_1(t_1), \dots, F_n(t_n))$, and continuity of the marginals makes those points the whole cube.

The converse needs no work at all. Given any copula C and any marginals, the composition $C(F_1(t_1), \dots, F_n(t_n))$ is a composition of a distribution function with non-decreasing right-continuous maps, so it is again a distribution function, and setting all but one argument to their upper limits returns F_i. This half is what does the damage in this chapter: it says the construction never fails, so no choice of C can ever be ruled out by the marginals.

When a marginal has an atom the argument breaks at exactly one point — F_i is no longer invertible, and (18.1) constrains C only on the range of F_i, which now has gaps. Existence survives, by interpolating across the gaps, and uniqueness does not. That is the whole content of the continuity hypothesis, and it is why a default-time distribution with a lump of probability at a coupon date is a case where "the" copula is not well defined. □

Remark (What it means)**.** A joint distribution factors cleanly into two independent pieces: the marginals, which say how each name behaves on its own, and the copula C, which says how they are coupled. The two can be chosen separately, and every choice is legitimate.

The intuition is a change of coordinates. $F_i(\tau_i)$ is uniform on $[0, 1]$ whatever τ_i was — this is the probability integral transform — so applying F_i to each name strips out everything specific to that name and leaves only its rank. The copula is the joint distribution of the ranks. It is what remains of the dependence after every marginal has been standardised away.

Sklar's theorem is usually presented as a technical result. For our purposes it is the precise statement of chapter 16's complaint: the market gives us the F_i and says nothing whatever about C, and by the converse half of the theorem, *every* C is consistent with the quotes. Choosing one is not calibration. It is an assumption, and it should be argued for rather than defaulted into.

Example 18.1 (The two extremes)**.** Two names, each defaulting within five years with probability 10%. What is the probability both do?

If they are independent, $C(u, v) = uv$ and the answer is 1%. If they are comonotone — one defaults exactly when the other does — then $C(u, v) = \min(u, v)$ and the answer is 10%. Both are consistent with identical CDS quotes on both names. A tenfold range in the joint probability, and nothing in the single-name market narrows it by a basis point.

18.2 Correlation Is the Wrong Word

The industry names this problem "correlation". That is a poor choice.

Remark (Correlation is not invariant)**.** Linear correlation is a property of the variables, not of their copula. Apply a strictly increasing function to one of them — take a log, or a square — and the copula is unchanged by construction, since ranks are unchanged, while the correlation moves. So correlation mixes up the dependence with the marginals, which is exactly the separation Theorem 18.1 was useful for making.

Rank statistics do not have this problem. Kendall's τ and Spearman's ρ depend only on the copula, and are the right things to quote if a single number must be quoted.

Theorem 18.2 (Attainable correlations)**.** *Let $X = e^{\sigma_1 Z_1}$ and $Y = e^{\sigma_2 Z_2}$ be lognormal. Then whatever the joint law of (Z_1, Z_2),*

$$\frac{e^{-\sigma_1\sigma_2} - 1}{\sqrt{(e^{\sigma_1^2} - 1)(e^{\sigma_2^2} - 1)}} \;\le\; Corr(X, Y) \;\le\; \frac{e^{\sigma_1\sigma_2} - 1}{\sqrt{(e^{\sigma_1^2} - 1)(e^{\sigma_2^2} - 1)}}. \tag{18.3}$$

Proof. The Hoeffding-Frechet bounds say the extreme joint distributions with given marginals are the comonotone one, $C = \min(u, v)$, and the countermonotone one, $C = \max(u + v - 1, 0)$. For lognormals these are $Z_2 = Z_1$ and $Z_2 = -Z_1$, and substituting each into $\mathrm{Corr}(e^{\sigma_1 Z_1}, e^{\sigma_2 Z_2})$ with the standard lognormal moments gives the two ends of (18.3). □

Example 18.2 (The bounds are not decorative)**.** Evaluating (18.3) numerically:[1]

Volatilities	**Lowest attainable**	**Highest attainable**
25% and 25%	−0.94	+1.00
100% and 100%	−0.37	+1.00
200% and 200%	−0.02	+1.00
50% and 200%	−0.16	+0.44

Read the last two rows. Two lognormals at 200% volatility cannot be more than two percent negatively correlated no matter how they are coupled, because the closest they can come to opposed is e^{2Z} against e^{-2Z}, and both of those are mostly near zero with an occasional enormous value — which makes them nearly independent, not opposed. And a 50% name and a 200% name cannot exceed 0.44 however tightly they are bound, because the fatter-tailed one has a variance the thinner one cannot track.

A risk system that accepts a correlation of −0.5 between two volatile positions has accepted a number that no joint distribution in the world can produce. It will not complain. It will return an answer.

Remark (And zero correlation is not independence)**.** The familiar one, included because it is the reason the next section exists. Correlation measures one particular linear summary of the joint law, and there are many ways to be strongly dependent with none of it. In particular, two variables can be uncorrelated and still almost always crash together, which is the case a portfolio of credits actually lives in.

18.3 The Number That Actually Matters

The failures above are diagnostic. This is the one that decides prices.

Definition 18.3 (Tail dependence)**.** The coefficient of upper tail dependence of a copula is

$$\lambda = \lim_{u \to 1^-} \mathbb{P}\big(U > u \mid V > u\big), \tag{18.4}$$

where U, V are the two uniform coordinates. It is the probability that one variable is extreme, given that the other is, in the limit of extreme.

[1] `quant/src/dependence.rs`.

Remark (Why this and not correlation)**.** Consider what a senior tranche is. It absorbs nothing until the portfolio has already lost, say, fifteen percent, which requires a large fraction of the names to have failed together. Its entire value is the probability of a joint extreme. It is a bet on λ and on almost nothing else.

Correlation, by contrast, is an average over the whole distribution, dominated by the middle where most of the probability sits. Two copulas can agree on correlation to three decimal places, agree on every marginal exactly, and disagree about λ completely — and a senior tranche will price off the disagreement rather than the agreement.

Theorem 18.4 (Tail dependence of the two standard copulas)**.** *The Gaussian copula with correlation $\rho < 1$ has $\lambda = 0$. The Student-t copula with correlation ρ and ν degrees of freedom has*

$$\lambda = 2\, t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right) > 0 \tag{18.5}$$

for every $\rho > -1$ and every finite ν.

Proof. Both halves come from the same question — conditional on one variable being far out in the tail, how far out is the other — and the two answers differ for one structural reason, which the calculation makes visible.

The Gaussian case. Write (X, Y) as standard bivariate normal with correlation ρ and condition on $X = x$. Then

$$Y = \rho x + \sqrt{1-\rho^2}\,\epsilon, \qquad \epsilon \sim N(0,1) \text{ independent of } X,$$

so λ is the limit of $\mathbb{P}(Y < x \mid X = x)$ as $x \to -\infty$, which is

$$\mathbb{P}\big(\epsilon < x\, \tfrac{1-\rho}{\sqrt{1-\rho^2}}\big) = \Phi\left(-|x|\sqrt{\frac{1-\rho}{1+\rho}}\right) \longrightarrow 0$$

for any $\rho < 1$, since the argument goes to $-\infty$. The conditioning moved Y's mean out by ρx, but the residual ϵ has a fixed scale, and a fixed-scale Gaussian cannot keep up with a barrier receding at rate x. So the second name is dragged part of the way into the tail and left there. Only at $\rho = 1$ does the drag become complete.

The Student case. The t copula is the Gaussian construction with one change: the pair is divided by an independent random scale,

$$(X, Y) = \sqrt{\frac{\nu}{W}}\,(Z_1, Z_2), \qquad W \sim \chi^2_\nu,$$

with (Z_1, Z_2) Gaussian of correlation ρ. Now conditioning on X being extreme is informative in a way it was not before: a large $|X|$ is evidence not only about Z_1 but about W being small, and a small W inflates Y as well. The tail event is explained by the common factor rather than by the individual one, and the common factor moves both names together.

Making that quantitative is a computation with the conditional law of Y given $X = x$, which for the bivariate t is again a t — with $\nu + 1$ degrees of freedom, a mean ρx, and a scale inflated by the factor $\sqrt{(\nu + x^2)/(\nu + 1)}$ that carries the information about W. Then

$$\mathbb{P}(Y < x \mid X = x) = t_{\nu+1}\left(\frac{(x - \rho x)\sqrt{\nu+1}}{\sqrt{(1-\rho^2)(\nu + x^2)}}\right),$$

and as $x \to -\infty$ the $\sqrt{\nu + x^2}$ in the denominator grows like $|x|$ and cancels the x in the numerator, leaving the finite limit

$$t_{\nu+1}\left(-\frac{(1-\rho)\sqrt{\nu+1}}{\sqrt{1-\rho^2}}\right) = t_{\nu+1}\left(-\sqrt{\frac{(\nu+1)(1-\rho)}{1+\rho}}\right).$$

The factor of two in (18.5) is the standard convention of adding the two tails. Note where the cancellation came from: the scale of the conditional distribution grew with $|x|$, and it grew because the mixing variable W is shared. That is the entire difference between the two copulas. □

Structure (The cancellation is the point)**.** The two calculations are the same calculation with one term changed, and the term is the scale of the conditional law. In the Gaussian case it is constant, so a barrier receding at rate $|x|$ eventually outruns it and the limit is zero. In the Student case it grows like $|x|$, the two rates match, and a finite limit survives.

The mechanism carries beyond copulas, and it is general: a fixed-scale residual cannot produce tail dependence, and a shared random scale always can. Any model in which the extreme behaviour of several quantities is driven by a common multiplicative factor — a stochastic volatility shared across names, a funding cost that hits every position, a liquidity parameter — has tail dependence for this reason, and any model whose only coupling is through the mean of a fixed-variance residual does not, whatever correlation is fed into it. Chapter 24's stress work is the practical form of the same observation: the scenario that matters is the one that moves the shared scale.

Remark (Read that again)**.** The Gaussian statement is not that tail dependence is small, or that it has been approximated away. It is exactly zero, for every correlation short of perfect. Condition on one name being in a one-in-a-million event and the probability that the other is too converges to zero. Joint extremes, in a Gaussian copula, do not happen.

The reason is the same one that makes a bivariate normal's contours ellipses: as you move out along the diagonal the density falls off faster than along either axis, so the far tail of a Gaussian is dominated by *one* variable being extreme, never both. The correlation controls how elliptical the contours are.

The Student-t fixes it with one extra parameter and no extra conceptual machinery. A t is a normal divided by an independent random scale, and it is that shared scale that produces joint extremes: occasionally the whole system is drawn from a wide distribution, and then everything moves at once. This is not a mathematical trick. It is a common volatility factor, and it is what the world does.

18.4 Four Things To Use Instead

Now the constructive half. These are ordered by how much structure they ask for, and the right choice depends on what is being priced and on whether it has to be hedged.

18.4.1 Give the copula tails

The cheapest fix, and often enough. Replace the Gaussian copula by a Student-t with the same correlation matrix and one degrees-of-freedom parameter. Everything about the implementation

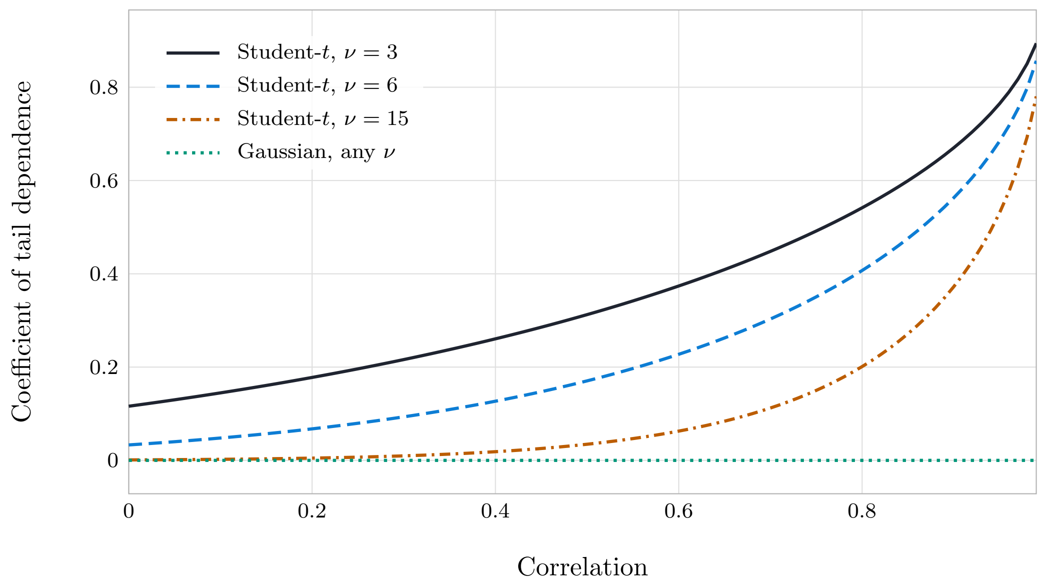

Figure 18.1: The tail dependence coefficient against correlation. The Gaussian curve is the horizontal axis — not close to it, on it, for every correlation up to one, where it jumps discontinuously to one. The Student-t curves sit well above and rise as the degrees of freedom fall. At a correlation of 0.3 a t with four degrees of freedom gives $\lambda = 0.16$: condition on one name being in an extreme event and there is a one in six chance the other is as well. The Gaussian model, calibrated to exactly the same correlation, says the chance is zero.

Drawn by `t_tail_dependence`.

is unchanged — the same factor structure, the same simulation, the same calibration to marginals — and λ goes from zero to (18.5).

The ν parameter can be calibrated to the tranche market, or set from history: fitting a t to equity index returns typically gives ν between three and six, and there is no reason credit should be thinner-tailed than equity.

Remark (Its limitation, stated honestly)**.** The t copula is symmetric: it produces joint booms as readily as joint crashes. Credit is not symmetric — names default together far more often than they are jointly upgraded — so a t that is fitted to the downside will overstate the upside. Whether that matters depends on whether anything in the book pays off on the upside. For a tranche, largely not.

18.4.2 Give it asymmetric tails

If the asymmetry matters, the Archimedean family provides it. A Clayton copula has lower tail dependence and none in the upper tail; a Gumbel copula the reverse. Each is a one-parameter family generated by a single convex function, which makes them easy to simulate and easy to reason about.

The cost is that Archimedean copulas in dimension n are exchangeable: they have one parameter for the whole basket and cannot express that two names in the same sector are more tightly coupled than two names in different ones. That is a serious limitation for a real portfolio, and the usual repair — nesting them hierarchically by sector — works but restores much of the complexity that made them attractive.

18.4.3 Model the cause, not the coupling

The first two are still copulas, and every copula shares a defect that the previous section did not mention because it is not about tails.

Remark (A copula has no dynamics)**.** A copula on default times is a static object. It gives the joint distribution of the τ_i as seen from today, and it says nothing about how that distribution evolves. There is no filtration in it and no process. So there is no notion of the model's own price tomorrow, and therefore no hedge ratio, and therefore no replication argument — which is to say that none of chapters 4 or 5 apply.

This is a deeper problem than the tails. A trader running a tranche book must hedge it with single-name CDS, and the hedge ratio has to come from somewhere. In a copula model it comes from bumping a static parameter and re-running, which is a finite difference of a quantity that was never claimed to be a price process.

The alternative is to model dependence where it comes from. Chapter 2 built default as the first jump of a Cox process with intensity λ_t; make the intensities share a factor,

$$\lambda_t^i = a_i\, Y_t + Z_t^i, \tag{18.6}$$

with Y a common process and Z^i idiosyncratic ones. Now everything is a process. Names default together because their intensities rise together; the joint distribution is an output rather than an assumption; and the model has a filtration, so a tranche has a genuine delta with respect to the single-name spreads that hedge it.

Remark (What this buys, and what it costs)**.** It buys dynamic consistency, which is the thing a copula cannot provide at any price. Spread moves and defaults are the same mechanism seen at different scales, which matches what a credit book actually experiences: correlated spread widening first, defaults later.

It costs calibration difficulty. Getting enough tail dependence out of (18.6) requires Y to be able to jump — a diffusive common intensity produces correlated spreads and still almost no joint defaults, which is the Gaussian problem in a new costume. So Y is usually taken with jumps, and the model is harder to fit than a copula. That is the trade, and for a book that has to be hedged rather than merely priced, it is usually worth making.

18.4.4 Do not model dependence at all

The fourth option is the most interesting, and it is available whenever a liquid tranche market exists.

Notice what a tranche actually is. It absorbs portfolio loss L between a and d, so its payoff is

$$\frac{(L-a)^+ - (L-d)^+}{d-a}, \tag{18.7}$$

which is a call spread on a single scalar variable. The whole capital structure is a strip of call spreads on L — and by the Breeden-Litzenberger theorem of chapter 9, a strip of call spreads on a variable determines that variable's distribution.

Remark (The same move as chapter 9)**.** This is the local volatility argument again. There we stopped trying to guess the dynamics of the underlying and read the risk-neutral density straight out of the option prices. Here we stop trying to guess the copula and read the loss distribution straight out of the tranche prices. In both cases the object that was being modelled turns out to be observable, and modelling it was never necessary.

The advantages transfer too, and so do the limitations. What is recovered is the loss distribution at each horizon and nothing more — the same partial information chapter 11 had to work around with local-stochastic volatility, and for the same reason. A bespoke tranche with different attachment points can be priced by interpolating the recovered distribution. A product depending on which names defaulted, or on the order they defaulted in, cannot: that is information about the copula and not about L, and it is not in the tranche prices.

Exercise. Show from (18.7) that the notional-weighted expected losses of any partition of $[0, 1-R]$ into tranches sum to the expected loss of the portfolio itself, whatever the dependence. (Hint: the call spreads telescope.) This is the credit analogue of put-call parity: a model-free constraint, which makes it the first thing to check in any implementation. `quant/src/dependence.rs` tests exactly this, for both copulas and three correlations.

18.5 The Gaussian Copula, and What It Cost

We can now say what went wrong precisely.

The Gaussian copula is not a broken model. It is a coherent, tractable, easily simulated dependence structure that reproduces any correlation matrix exactly, and for a great many purposes it is perfectly adequate. Its one structural property is Theorem 18.4: it sets tail dependence to zero.

It was then adopted as the market standard for pricing tranches, whose senior end is a pure bet on tail dependence. The single quantity those instruments were sensitive to was the single quantity the model asserted to be zero.

Example 18.3 (What the assumption is worth)**.** A large homogeneous portfolio, five percent of names defaulting by the horizon, forty percent recovery, thirty percent correlation. The same marginals, the same correlation, under a Gaussian copula and a Student-t with four degrees of freedom:[2]

Tranche	**Gaussian**	**Student-t**	**Ratio**
$0-3\%$ (equity)	54.11%	39.40%	0.73
$3-7\%$	19.58%	18.25%	0.93
$7-15\%$	5.83%	8.35%	1.43
$15-30\%$	0.81%	2.41%	2.98
$30-60\%$ (super senior)	0.019%	0.195%	10.16

Every number a risk report records is identical between the two columns: the same default probabilities, the same recovery, the same correlation. The super senior tranche is worth ten times more under one than the other.

And note the direction. The fat-tailed copula makes the equity tranche *safer* — it puts more probability at low losses as well as at high ones. So the effect is not a uniform increase in risk that a conservative haircut would have caught. It is a transfer of risk up the capital structure, and it is invisible at the bottom, where the market's attention was, and largest at the top, where the notional was.

Remark (Base correlation was the tell)**.** The market did notice, in the way markets do. Fitting one Gaussian correlation to all tranches simultaneously turned out to be impossible: each tranche implied a different one, and the market began quoting a "base correlation" per attachment point, rising steeply with seniority.

This is precisely chapter 9's volatility smile, and it means precisely the same thing. A parameter that the model says is a single number, quoted instead as a curve, is the market saying the model is wrong in a structured way and telling you the shape of the error. In the volatility case, the profession read the message and built chapters 10 to 12 in response. In the correlation case, the curve was largely treated as a quoting convention.

The lesson generalises past credit. When a model requires a different parameter value for each instrument, the parameter is not a parameter. It is a record of the model's error.

Remark (The fair verdict)**.** It is fashionable to blame the formula, and that is too easy — and it lets the actual mistake escape unexamined. The Gaussian copula did what it says. The failure was in the layer above it: an instrument was priced with a model that had been chosen for tractability and for its fit to the part of the capital structure that traded, and nobody asked what the senior end was sensitive to and whether the model had anything to say about it.

That question — what is this instrument's payoff a bet on, and does my model have a view on that quantity or has it assumed one — is not specific to credit, and it is the one thing to take away from this chapter.

[2] every tranche in the table.

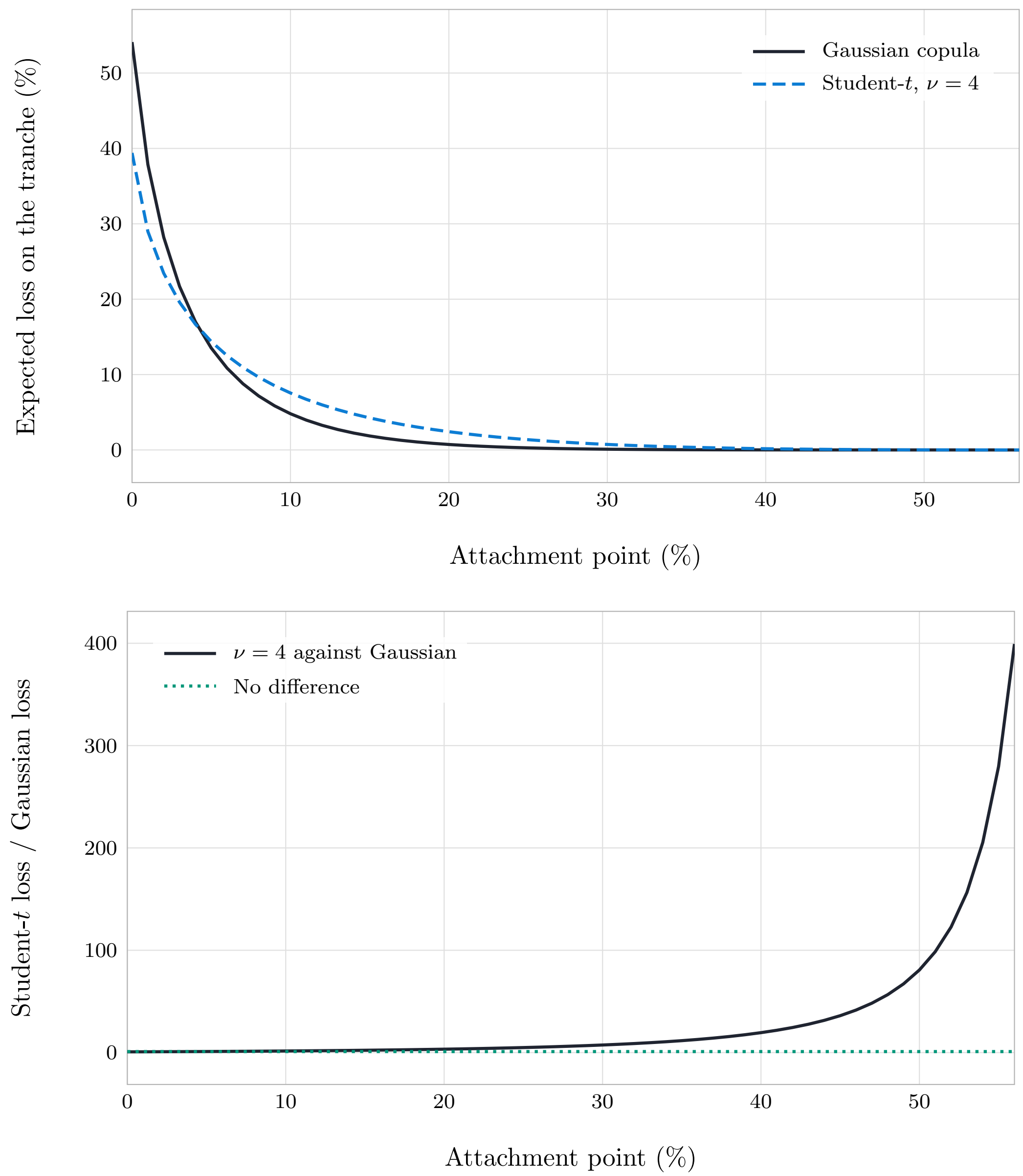

Figure 18.2: The same portfolio sliced into three point wide tranches at every attachment point. The upper panel shows the levels, which are dominated by the equity end and reveal nothing about the top. The lower panel shows the ratio of the two models, which reveals everything: below one at the bottom, crossing at around six percent attachment, and running away above it. A model chosen for its behaviour on the tranches that trade most was applied to the tranches that carried the most notional.

Drawn by `Portfolio::tranche_expected_loss` and `Portfolio::exceedance`.

18.6 The Same Gap in Rates

A CMS spread option pays on the difference between two swap rates of different tenors,

$$\big(S_1(T) - S_2(T) - K\big)^+, \tag{18.8}$$

and the callable range accruals and steepener notes that make up much of the structured rates market are strips of them. Chapter 12 met this payoff already, as the instrument that exposes what a one-factor model cannot do.

Now notice the shape of the pricing problem. Each rate's marginal distribution is *known*: chapter 15 obtains it from that rate's own swaption smile, by replication, without any model of the curve. What is not known is the joint distribution, and (18.8) depends on nothing else. That is exactly the position §18.1 opened this chapter with, in a different market: marginals supplied by the market, dependence supplied by the modeller.

So the market does exactly what credit did. It takes the two marginals from the smiles, joins them with a Gaussian copula carrying a single correlation, and integrates (18.8) against the result. And it has arrived at the same symptom: no single correlation fits the quoted spread options across strikes, so a correlation is quoted per strike — the base correlation of the previous section, in rates, for the same reason.

Calculation 18.5 (What the copula is worth on a spread option)**.** Set up as the tranche comparison was, so that nothing but the dependence differs. Two rates at 4% and 2.5%, both with an absolute volatility of 100 basis points, five years, correlation 0.8 — and both keeping exactly the same normal marginal under either copula.[3]

Against a Student-t with four degrees of freedom, the option struck at a spread of 1.5% is worth 0.95 times its Gaussian value. Struck at 4.5% it is worth 2.6 times.

Tail dependence makes extremes arrive together, and one expects two rates that move together to produce a narrower spread. What actually happens is that a Student-t copula is a bivariate normal divided by a common random scale, and a small draw of that scale makes both rates extreme *without making them equal*. The spread inherits the heavy tail. So the Gaussian copula understates a far strike spread option for precisely the reason it understated a senior tranche, and the mass it is missing there is the mass it has put in the middle.

Remark (Two assumptions, not one)**.** Rates adds a layer that credit does not have.

The marginals in credit come from CDS quotes with no modelling in between. The marginals here come from chapter 15, which reaches them only through the linear annuity mapping — an assumption in its own right, and one calibrated to different instruments than the copula is. A CMS spread option therefore rests on two dependence-free approximations stacked on each other: a mapping to get each rate's law under the payment measure, and a copula to join them. Neither is implied by the other and a reserve against one says nothing about the other.

A quanto CMS spread adds a third, since chapter 17's adjustment needs each rate's correlation with the exchange rate. The object required is then a three-dimensional joint law of which the market quotes only pairwise pieces, and the standard practice — assemble it from those pieces and hope the assembly is consistent — is not guaranteed to produce a valid joint distribution at all.

[3] `measured`. The Gaussian column is checked against Bachelier, which it must reproduce, since normal marginals joined by a Gaussian copula give a normal spread.

Which leaves the choice this book keeps returning to. A term structure model of chapter 12 or chapter 13 supplies the joint law by construction and gives up the exact fit to the smiles; a copula keeps the exact fit to each marginal and supplies the joint law by assumption. Chapter 12 measured what the first costs on this very product. Calculation 18.5 measures what the second costs.

References

- Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Publications de l'Institut de Statistique de l'Universite de Paris*, 8, 229–231.

- Embrechts, P., McNeil, A., & Straumann, D. (2002). Correlation and dependence in risk management: properties and pitfalls. In *Risk Management: Value at Risk and Beyond*, Cambridge University Press, 176–223.

- Li, D. X. (2000). On default correlation: a copula function approach. *Journal of Fixed Income*, 9(4), 43–54.

- Duffie, D., & Garleanu, N. (2001). Risk and valuation of collateralized debt obligations. *Financial Analysts Journal*, 57(1), 41–59.

- Frechet, M. (1951). Sur les tableaux de correlation dont les marges sont donnees. *Annales de l'Universite de Lyon, Section A*, 14, 53–77.

- Hoeffding, W. (1940). Massstabinvariante Korrelationstheorie. *Schriften des Mathematischen Instituts der Universitat Berlin*, 5, 179–233.

- Nelsen, R. B. (2006). *An Introduction to Copulas*, 2nd ed. Springer.

- McNeil, A. J., Frey, R., & Embrechts, P. (2015). *Quantitative Risk Management*, revised ed. Princeton University Press.