Global Derivatives, Paris, 2011

Model Risk Management for Interest Rates, Funding and Credit

Massimo Morini

IMI Bank of Intesa San Paolo, Head of Credit Models and Coordinator of Model Research*
Milan Bocconi University, Professor of Fixed Income

*This presentation expresses the views of its authors and does not represent the opinion of Banca IMI, which is not responsible for any use which may be made of its contents.

Outline

- Different views on Model Risk
- Model Uncertainty
- Accounting, Fair Value and Model Risk
- Regulators on Model Risk after the Crisis
- A practical scheme for Model Validation
- Comparing Models
- Example on Gap Risk with Reduced-Form vs Structural Models
- Hints at Local vs Stoch Vol Models

Outline

- When market changes model: example on interest rates in the credit crunch
- Model evolution and foundations: example on funding liquidity in pricing
- Validating approximations vs montecarlo
- The swaption formula in LMM and the CMS convexity adjustment
- Validating Approximations vs Analytics. Example on SABR
- Real Hedging Test. Local vs Stochastic Volatility

Outline

- Stress Testing Models and Stress Testing with Models
- Portfolio stress testing. Example on Correlation Skew
- Pitfalls in stress testing. Example on Dynamic Copulas
- Model stress testing. Example on Mapping with Comparables
- When the payoff is wrong
- Example on Bilateral Counterparty Risk

Different views on Model Risk

A Model Validation practice requires agreeing on what we mean by Model Risk, from a practical perspective.

There are two main approaches to the problem. One, that we call Value approach, can be described using the words of Derman (1996, 2001). The other, that we call Price approach, is described in Rebonato (2003).

The Value approach

The model must provide *"a realistic (or at least plausible) description of the factors that affect the derivative's value"*

In this sense **model risk is the risk that the model is not a realistic/plausible representation** of the reality, or better still **of that part of reality that affects in a relevant way the derivative value**. Thus the errors a modeler makes can be:

"You may have not taken into account all the factors that affect valuation. You may have incorrectly assumed certain stochastic variables can be approximated as deterministic. You may have assumed incorrect dynamics. You may have made incorrect assumptions about relationships".

Derman E., GS. *Model Risk*.

The Value approach. Which prescriptions?

The points of a Value approach to Model Validation, taken from Derman (2001), are given by the solution to these questions

1. Is the payoff accurately described?
2. Is the software reliable?
3. Has the model been appropriately calibrated to the prices of the simpler, liquid constituents that comprise the derivative?
4. Does the model provides a realistic (or at least plausible) description of the factors that affect the derivative's value?

The heaviest assessment is the last one, that requires, following Derman (1996), the analysis of such things as dynamics, number of factors, relations... It must be however clear that "*a model is always an attempted simplification*" of a reality, and as such there can be no perfectly realistic model.

The Price approach

"Model risk is the risk of occurrence of a significant difference between the mark-to-model value of a complex and/or illiquid instrument, and the price at which the same instrument is revealed to have traded in the market".

Rebonato R., RBS, *Theory and Practice of Model Risk Management*.

This view points out that real losses that hit an institution's balance sheet usually do not appear "*because of a discrepancy between the model value and the 'true' value of an instrument*", but through the mark-to-market process, because of a discrepancy between the **model value and the market price of an instrument**. **As long as the market agrees with our valuation, we do not have large losses due to models**, even if *"market prices might be unreasonable, counterintuitive, perhaps even arbitrageable"*. We have no mark-to-market losses, we can sell at the value at which we booked.

The Price approach. Which prescriptions?

Let'see. This can happen for three reason:

1. The reason can be that we were using a model different from the market consensus. From this comes the first prescription of the price approach, given strongly in Rebonato (2003), to gather as much information as possible on the approach used by the market players. This requires observing as many trades as possible, look at collateral regulations, exploit market consensus platforms, in order to make reverse engineering of the observed prices.
2. If we use the same model prevailing in the market and this model is not changing, the only cause can be operational errors (legal, software...), that should be avoided.
3. Otherwise, we can have losses because the market consensus suddenly changes (or a market consensus did not exists and it suddenly emerges). From this comes the prescription to *"surmise how today's accepted pricing methodology might change in the future"* (Rebonato (2003)).

Comparing the approaches

Are Value and Price approach so different in practice?

- Both require controlling the operational risks associated to the use of models (implementation and payoff description).
- The Price approaches focuses on collecting information about the market consensus, and points out for the first time the importance of this aspect. This is not however in contrast with the Value approach that requires calibration to liquid market prices and includes assessment of "*market sentiment*" among the factors that affect a derivative value.
- Then the Price approach requires to analyze how market can change, to be ready to such an event. This happened recently for mortgage-backed securities after the burst of the subprime crisis, when the rating-sustained approach to price tranches with gaussian copula plus deterministic spreads and historical (or loosely mapped) correlations and recoveries broke down.

When the Market consensus changes

How can the market suddenly change the model consensus, as it happened in the subprime crisis? Often because new market events related to fundamentals (default clusters in the subprime markets, decreasing house prices, the abnormal downwards movement of equity prices in 1987 that led to the appearance of the smile) or new pieces of research show that some of the assumptions of the model consensus were unrealistic in a relevant way (Rebonato quotes the paper 'How to throw away a billion dollars' on Bermudans that contributed to moving from unrealistic one-factor models to multifactor models for interest rate derivatives).

Also in the price approach the correct specification of the model, or its realism, have to be taken into account to prevent model losses, since lack of consistency with reality is the main fact that can lead to sudden changes in model consensus.

A synthetic view

- Knowledge of the market approaches and analysis of model correct specification remain crucial for both approaches.
- Understanding the market approach requires, as Rebonato (2003) says, a battery of models to perform reverse engineering. Which model can be consistent with a given price? Assessing that the market is not incorrectly specified requires comparing different models. Which one is more consistent with real behaviour, with the unavoidable approximations? Model comparison is crucial for both.
- In both cases we want to understand clearly the implications of the application of a model to a given payoff. Guessing how a model can change, and balance between market sentiment and realism require understanding the model deeply.

Model Misunderstanding is the ultimate Model Risk for both approaches.

A scheme for Validation

We show in the following an example of a model analysis based on these points

1. Calibration: the models need to fit all relevant market information.
2. Reasonableness: rule out those model assumptions who are obviously unrealistic on relevant features.
3. Market intelligence: investigate what approach is used by other players.
4. Reality check: study the behaviour of the underlying to understand which models are more realistic.

These make sense if applied in model comparison, with a battery of models:

A Detecting, among the models selected above, the two extreme opposites in terms of payoff value. This is useful for judging on model realism and understanding which models are used by the other market players, via reverse engineering.

B Constructing a family of intermediate (parametric) models to make a choice, quantify the residual model risk, and have a tool for adopting the appropriate model risk management

Foundations of Modelling

The price of a derivative that pays X_T at time T is usually written as

$$\Pi_t = \mathbb{E}^Q \left[D(t,T) \, X_T \middle| \, \mathcal{F}_t \right] \tag{1}$$

There are a number of symbols to describe here. First, $D(t,T)$, the discount factor:

$$D(t,T) = e^{-\int_t^T r(s)\,ds} = \frac{B(t)}{B(T)}.$$

The quantity $B(t)$ is the value at t of the *money market account,* the mathematical representation of a bank account, growing at a rate of interest $r(t)$ according to

$$dB(t) = r(t) \, B(t) \, dt \tag{2}$$

so that, by solving the very simple *deterministic* equation (2), we find out that if we invest 1 at time 0 we have at t

$$B(t) = e^{\int_0^t r(s)\,ds.}.$$

Foundations of Modelling

$$\Pi_t = \mathbb{E}^Q \left[D(t,T)\, X_T \middle| \mathcal{F}_t \right]$$

The symbol

$$\mathbb{E}^Q \left[\cdot \middle| \mathcal{F}_t \right]$$

indicates *expectation*, using the available *information at* t, indicated as $\mathcal{F}_t$, under a *risk-neutral* or *risk-adjusted probability measure* Q.

Usually the formal framework is a triplet called a probability space

$$(\Omega, \mathcal{F}, Q)$$

Valuation Uncertainty

The analysis of the alternative realistic model and the understanding of the models used in the market does not eliminate model uncertainty. How should we deal with this residual uncertainty? Prices are supposed to be expectation of uncertain quantities...shouldn'we take expectation of all that is uncertain and this is the price, with no uncertainty left? Not necessarily. Think of a model with random parameters (Brigo and Mercurio 2003). We know the underlying of an option is a diffusion

$$dS\left(t\right)=rS\left(t\right)dt+\sigma^{I}S\left(t\right)dW\left(t\right)$$

but we have uncertainty on the value of volatility, in the sense that

$$\sigma^{I}=\begin{cases}\sigma^{1}\text{ with prob }p_{1}\\ \sigma^{2}\text{ with prob }p_{2}\end{cases}.$$

Scenarios depend on random variable I, drawn at $t=\varepsilon$ infinitesimal after 0, independent of W and which takes values in $\{1,2\}$ with probability

$$Prob\left(I=i\right)=p_{i}.$$

A model with random parameters vs uncertainty on the parameters

How do we price the option with strike K and maturity T? We have as usual that the option price is

$$\Pi(K,0,T) = P(0,T)\,\mathbb{E}\left[(S(T)-K)^{+}\right].$$

We can use the law of iterated expectations and write

$$\Pi(K,0,T) = P(0,T)\,\mathbb{E}\left[\mathbb{E}\left[(S(T)-K)^{+}\,|\sigma^{I}\right]\right].$$

Notice that the external expectation regards a random variable that can take only discrete values, so the expectation is given by a weighted average of these values, where the weights are the probabilities of the different volatility scenarios:

$$\Pi(K,0,T) = P(0,T)\sum_{i=1}^{2} p_i\mathbb{E}\left[(S(T)-K)^{+}\,|\sigma^{I}(t)=\sigma^{i}(t)\right].$$

Now look at the inner expectation. Taking into account that $\sigma^{I}(t)$ is independent of $dW(t)$, we have that, *conditional to* $\sigma^{I}(t)=\sigma^{i}(t)$, $S(t)$ is just a geometric brownian motion with volatility $\sigma^{i}(t)$.

A model with random parameters vs uncertainty on the parameters

Therefore the option price is just the average of two Black and Scholes price.

$$p_1 BS\left(S\left(0\right), K, T, \sigma^1\right) + p_2 BS\left(S\left(0\right), K, T, \sigma^2\right) \tag{3}$$

Here there is no model uncertainty, the model is only one, and therefore there is no valuation uncertainty. Notice that if we had to perform a simulation, for each scenario we should first - at an infinitesimal time ε - simulate σ^I, then once we know that on this scenario $\sigma = \sigma^i$, with $i = 1$ or $i = 2$, we use this volatility until maturity. This is *one specific model* with random volatility, it is not a representation of model uncertainty.

A model with random parameters vs uncertainty on the parameters

Price uncertainty arises when we have model uncertainty. Model uncertainty would be if: we have the same uncertainty on the future volatility that we had in the above model, but

1. at ε we will not know the right volatility
2. we have no idea of what p_1 and p_2 are
3. other market players may know what is the real value of σ, or may have even less information than us.

This makes the picture different. How to deal with this?

Model uncertainty

Cont (2006) deals with such issues and says that in reality market operators work with a multiplicity of models, or triplets

$$
\begin{aligned}
&(\Omega, \mathcal{F}, Q_i)\,, \\
Q_i \quad &= \quad Q_1, Q_2, \ldots Q_N
\end{aligned}
$$

He calls uncertainty on the right probability measure - the right model - a type of *knightian uncertainty*, from Knight (1921). Here we are unable to give a distribution to the future events, as opposed to cases of uncertainty where we do not know the future outcome but we know its probability distribution (the roulette!). The latter is what we usually call ‘risk’.

The bayesian approach

Hoeting (1999) suggests to use, in a bayesian framework, an approach similar to the averaging of formula (3) also for model uncertainty. But this makes no sense in our context, since it makes uncertainty *on* the model undistinguishable by uncertainty *in* the model.

Standard models that want to be realistic treat uncertain parameters as observable stochastic processes. In the latter case uncertainty is resolved by the passage of time. But model uncertainty has the specific feature that does not necessarily get revealed progressively by the sheer passage of time, it does not regard the future evolution of observable stochastic processes but today's distributions.

Worst-case approach dignified

Cont (2003) notices that the typical approach of banks when faced with model uncertainty is not to average across models but to adopt a worst case approach. If we were option buyers and $\sigma^2 > \sigma^1$, we would adopt σ^1. There is one way to be fully consistent with the worst-case principle, a conservative way to make sure we will not suffer losses from any mispricing:

$$P^{ask} = BS\left(S(0), K, T, \sigma^2\right), \; P^{bid} = BS\left(S(0), K, T, \sigma^1\right).$$

Gibdoa and Schmeider (1989) show that one can even formalize the old 'conservative approach' used so often in banks. The worst case approach corresponds to maximization of utility when we are faced with total ignorance on the probabilities.

Measures of Model Uncertainty

Following this line, Cont (2003) proposes two measures of model uncertainty: one is akin to

$$\max_{Q_{i=1,\ldots,N}} \text{Price} - \min_{Q_{i=1,\ldots,N}} \text{Price},$$

the second one weights more or less models depending on the higher or lower capability to price liquid market instruments. For Cont there is no model uncertainty when:

1. the market is liquid (model uncertainty is always lower than bid-ask spread!)
2. we can set up a model-independent static hedge

Later on we discuss this. We show a case where model risk stroke a liquid market... and this market became illiquid. And we will see how a static hedge that appears model-free can suddenly break down because the market consensus on the model changes.

Market Completeness and Uncertainty

How does this relate with the theory of incomplete markets? Formally, the above market would be incomplete if the underlying $S(t)$ was not a traded asset. In this case (let us now make the parameters time-dependent) we should price assuming

$$dS(t) = (\mu_t - \sigma_t \gamma_t) S(t) dt + \sigma_t S(t) dW(t)$$

where γ is my price of risk. The drift is not necessarily equal to r since the underlying is not traded so there are no 'no-arbitrage conditions' that force this equality. Thus, even if all market players knew the parameters under the real-world probability measure, they would not know the drift under the measure that matters for pricing, and there would be model uncertainty.

Market Completeness and Uncertainty

But this representation, although fully consistent with the theory, clashes with reality. If the underlying is an illiquid financial asset, we will be most often in a situation where we know neither

$$\mu_t, \gamma_t, \sigma_t.$$

The uncertainty on γ is explained by the standard theory and will be resolved looking at my risk aversion. The uncertainty on μ and σ is not explained by the standard theory (since the standard no-arbitrage theory assumes all players know and use the same model) but is there in the practice. It will also be solved looking at risk aversion...

In the final estimation of the drift

$$\mu - \sigma_t \gamma$$

the effect of my arbitrary risk aversion on γ and the effect of my uncertainty about the real-world parameters μ and σ_t will be indistinguishable.

Market Completeness and Uncertainty

Thus, in the practice of financial markets, the problem of market incompleteness and the problem of model uncertainty interact strongly. Uncertainty on the risk-neutral adjustment $Q|P$ compounds with uncertainty on the P dynamics under the real world measure. But there is more. If we introduce jumps or stochastic volatility, we would need as many traded instruments as independent risk factors to make the market complete. The odds are that real markets are incomplete almost by definition; so on many of their parameters risk aversion compound with model uncertainty.

Accounting for Modellers

The fair value principles prescribed by the international accounting standards have a strong influence on how pricing models must be designed and applied. Here we revise and analyze these principles and their link to model risk, including the revisions or specifications that were introduced after the recent credit crisis.

Then we review the classification of prices based on accounting principles (level 1,2 and 3) and show how these levels can be put in relation with different amounts of model risk, but also how different these concepts can be in many situations. This is explained with an example: the valuation of interest rate swaps - usually considered simple and liquid products - involves the use of comparables, extrapolation, and is based on strong assumptions on the nature of liquidity and credit risk in the interbank market. This is confirmed by the changes in swap valuation currently happening as a consequence of the credit crunch.

Implications of ACCOUNTING STANDARDS on MODEL RISK

Models are used every day along the life of a derivative, from inception to maturity, to update its value for accountancy purposes. It is in this accountancy process, and not in real money, that most profits and losses emerge for banks. The fundamental principles given by national or international accountancy boards have something to do with model risk.

The principles that must be followed for the valuation of financial assets and liabilities are given by organizations such as the International Accounting Standards Board (IASB) that has published the International Accounting Standards (IAS) and the International Financial Reporting Standards (IFRS). Then in all countries there are autonomous local bodies; the most important example is the FASB, that formulates accounting principles (GAAP) valid for the US.

FAIR VALUE

According to both IASB and FASB, derivatives and all financial assets or liabilities which are not held for an investment until maturity (banking book products) but can be bought and sold (trading book products) *should be evaluated following the principle of fair value*.

The IAS 39 document, issued first in 1998 and regularly revised, gives the IASB definition of fair value as

"*the amount for which an asset could be exchanged, or a liability settled, between knowledgeable, willing parties in an arm's length transaction*" (IAS 39)

while the FAS 157 document issued in September 2006 from FASB states that

"*Fair value is the price that would be received to sell an asset or paid to transfer a liability in an orderly transaction*" (FAS 157)

FAIR VALUE

Let us point out a few relevant and non-trivial implications of the above definitions.

1. The market is the fundamental reference to give a value to a derivative (but consider also point 4...).
2. Fair value is not a mid-price. In case of an asset it must be the sell price, for a liability we must consider the price required to extinguish it. It is an *exit value*.
3. Everything that would enter the sell-price must be considered, including the default risk of our counterparty (*credit value adjustment, CVA*) and, possibly, our default risk (*debt value adjustment, DVA*).
4. The expression *"arm's length transaction"* or *"orderly transaction"* mean that we must not consider a sale where one counterparty has a power on the other or when we are in a distressed market. The transaction must happen under normal market conditions.

FAIR VALUE and the CRISIS

Thus the accountancy view appears more in line with the Price approach to model risk. But it still raises some questions. The main one is: what should we do if we have reasons to think the market price is wrong? This argument was raised often during the credit crunch. The accountancy boards answered on September 30, 2008, when SEC and FASB issued a joint clarification, where they pointed out that fair value definition speaks of market prices in *"orderly transactions"*. When a market is affected by illiquidity and panic there cannot be orderly transactions, thus market prices may not constitute fair value - and thus can be corrected by personal judgment. What do they mean by correct prices? Correct prices need to come out of *reasonable risk-adjusted, liquidity-adjusted and credit-adjusted "expected cash flows"*. See "SEC Office of the chief accountants and FASB Staff Clarification on Fair Value Accounting".

FAIR VALUE and the CRISIS

Just a fortnight later the IASB issued a similar paper, where they claim that a firm should not "*automatically conclude that any observed price is determinative of fair value*", and in case they think it is not they should make "*adjustments to the transactions price that are necessary to measure the fair value*".

These 'clarifications' claim that fair prices should be descriptions of expected future cashflows, and need to coincide with market prices only when market prices can be broadly believed to be already such descriptions. The mark-to-market principle can be mitigated by considerations that regard the reasonablenss of market prices, at least when markets are judged to behave in a clearly irrational way.

This mirrors the debate we have seen between Value and Price approaches.

Level 1,2 and 3

There is another similarity between the accountancy prescriptions and the principles of Model Risk, in particular the link between model risk and lack of liquid market information. There exists in accounting principles a three-level hierarchy that classifies the prices based on the inputs that enter the valuation procedure; prices based only on objective inputs - coming from liquid markets and considered surely appropriate for the product to price - rank 1, while prices based on inputs computed making use of a plainly high amount of personal judgement rank 3. We present this hierarchy, although in the end we show by one examples why it must be taken with a pinch of salt if we want to perform sound model risk management.

Level 1,2 and 3

1 Level 1 (Liquid specific quotes): the price comes from a *liquid* market where products *identical* to the derivative to price are traded.

2 Level 2 (Comparable quotes or Illiquid quotes). These fair value prices come from one of the following two cases:

- a *liquid* market where products *similar* (not identical) to the derivative to evaluate are traded.

- a market where products *identical* to the derivative to evaluate are traded but the market is *illiquid*.

3 Level 3 (Model with non-quoted parameters). These fair value prices come from the use of a valuation technique that requires inputs which involve a crucial amount of personal judgement of the institution computing the price. This kind of prices are often called 'marked-to-model'.

Level 1,2 ,3 and Model Risk

According to the general interpretation, model risk should regard surely level 3 prices, in some cases also level 2 prices, and not at all level 1 prices. These should only involve 'market risk' and not 'model risk'. Formally this makes sense, but for the purposes of model risk management it must be taken with a pinch of salt. Level 1 prices are those taken from liquid prices of products *identical* to the derivative under analysis. A Level 1 price can involve model risk because it hinges on the subjective decisions that two assets are *identical*, while this may be true only under some hidden assumption that can suddenly become unrealistic.

For pointing out how model assumptions enter in the evaluation of the simplest vanilla derivatives we show the model assumptions hidden in the valuation of a simple swap. By the way, after the credit crunch these model assumptions are not valid anymore and what could be considered comparables in the market are not anymore (basis spreads for product with different tenors), which is leading to a global reform of swap pricing. The topic will be examined in detail later on.

The simplest swap

In the interbank market we can observe quotes for loans from bank to bank. Banks will agree on the rate at which such loans must happen: this rate has been called Libor for many years. For the maturity T, the market prevailing rate is $L(t, T)$. This means that if the notional of the loan is $1\$$, the lender gives 1 at t, and the borrower must give back

$$1 + L(t, T)(T - t)$$

at T. But also... if the loan has a notional of

$$\frac{1}{1 + L(t, T)(T - t)}$$

the lender gives $\frac{1}{1+L(t,T)(T-t)}$ at t, and the borrower must give back 1 at T. Thus we see that this market, at least for the maturity T, gives directly the spot rate and also the value of a contract that pays 1 at T. The rates that we see quoted in the market at t are $L(t, \alpha)$ and $L(t, 2\alpha)$.

The simplest swap

Now we are a bank that has to price, at time $t = 0$, a Forward Rate Agreement with fixing at α and payment at 2α, the simplest one-period swap where at 2α one party has a payment indexed to a fixed rate K and the counterparty instead pays a floating rate that fixed at α:

$$(L(\alpha, 2\alpha) - K)\,\alpha. \tag{4}$$

We can trivially price it as follows. Lend $\frac{1}{1+L(0,\alpha)\alpha}$ to your counterparty for maturity α. You will receive 1 at α. Now lend her again this 1 for a maturity α, namely from α to 2α, and you will receive at 2α

$$1 + L(\alpha, 2\alpha)\,\alpha$$

At 0 one must also borrow $\frac{1+K\alpha}{1+L(0,2\alpha)\alpha}$ from the counterparty until 2α. We will pay at 2α

$$-1 - K\alpha$$

Put together with what we receive and get

$$L(\alpha, 2\alpha)\,\alpha - K\alpha,$$

The simplest swap

$$L(\alpha, 2\alpha)\,\alpha - K\alpha,$$

This is exactly the same payoff from the FRA given in (4).

Thus the price of the FRA must be equal, in an arbitrage-free market, to the price of the replication strategy:

$$\frac{1}{1 + L(0,\alpha)\,\alpha} - \frac{1 + K\alpha}{1 + L(0,2\alpha)\,\alpha}.$$

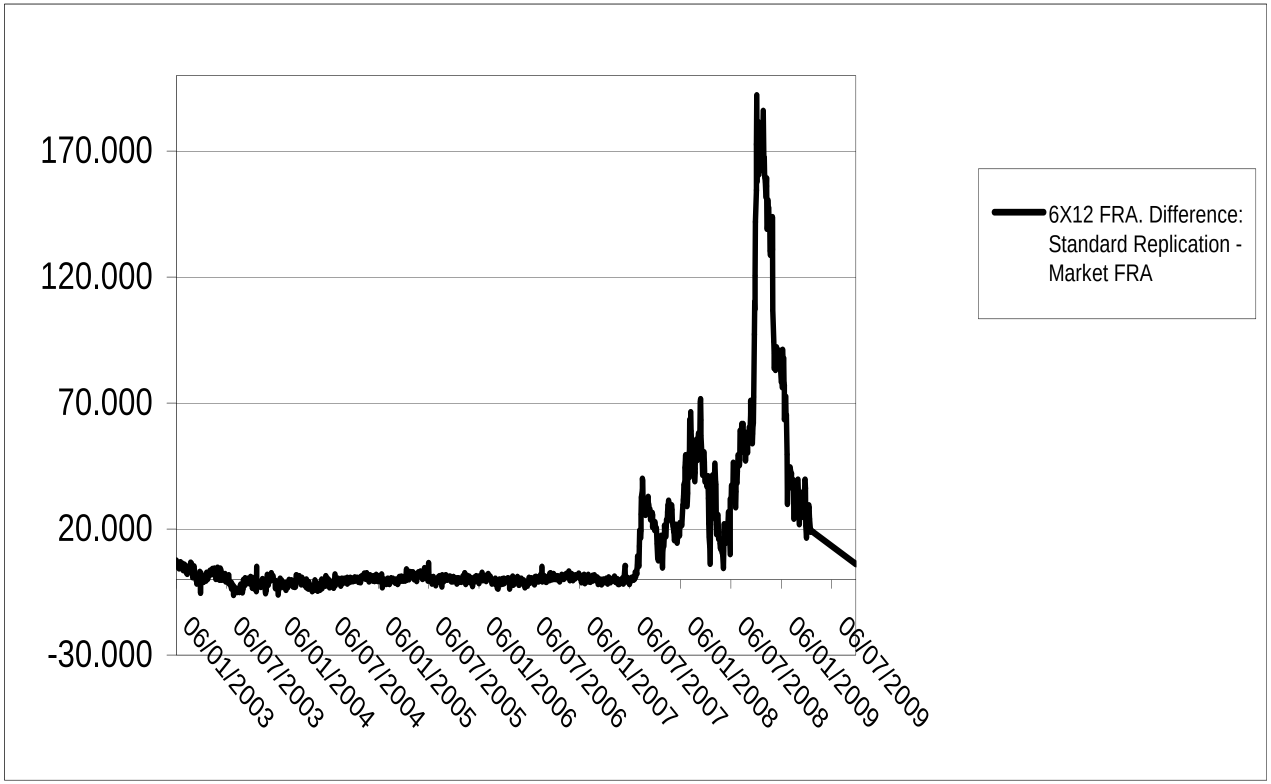
The simplest swap and Model Risk

Is this pricing level 1, 2 or 3? We are working on liquid markets, we are not using any comparable, but only things that ought to be the same in any arbitrage-free market, and we are using no evaluation model. It looks like level 1. Usually a FRA is not level 1 but level 2 because the evaluation goes through a bootstrap of the term structure, that always involves at least interpolation between maturities, but here we assumed to have rates and thus discounts for exactly the maturities of the FRA. The FRA appears *identical* to the portfolio of loans; the payoffs are the same.

And yet this identity comes from some strong model assumptions, that go far beyond the assumption of no arbitrage, and as such it is subject to model risk since the underlying model assumptions, like all model assumptions, can be rejected by the market in the future. And they were rejected by the market in one single day, August 9 2007.

2007: the replication breaks down

Below we report the difference between the FRA price and the value of the spot replication for a period of more than 6 years.



The simplest swap and Model Risk

Where are the hidden model assumptions? First of all, the FRA above is treated as default-free. What happens when we introduce risk of default in the interbank market, that exploded with the subprime crisis?. We will see that this does not necessarily break the replication because the payoff and the replication strategy may be affected by risk of default in the same way. The important effect of default risk is that *our counterparty in the replication may not have the same risk of default expressed by the rate* $L(\alpha, 2\alpha)$ *that appears in the payoff*. My counterparty C will have a fair lending rate $L_C(\alpha, 2\alpha)$ which is not necessarily the same as the rate $L(\alpha, 2\alpha)$, and this breaks a crucial fact used in the replication. This displays the imprecision of the above payoff description: what does it mean that the FRA will play a floating rate $L(\alpha, 2\alpha)$ at 2α? What is that rate, precisely, how will it be computed at 2α? The rate $L(\alpha, 2\alpha)$ is usually a Libor rate that comes out of a rather elaborate process, as the average of a Libor panel.

The simplest swap and Model Risk

We will see that we can consider the replication at least approximately valid only if all banks in the panel and also my counterparty have very similar default risk and *this will be expressed at* α *by the rate*

$$L(\alpha, 2\alpha) = L_C(\alpha, 2\alpha).$$

But this hangs on the following assumptions:

1. All banks active in the market have a negligible risk of default
2. If this risk of default is not negligible, it is the same for all banks and its volatility is negligible
3. If this volatility is not negligible, in any case all banks involved have risk of default perfectly correlated

These conditions were approximately valid for the ten years preceding the credit crunch, but in the summer of 2007 they all fell down, from the first to the last.

The simplest swap and Model Risk

This emerged formally as market risk for quoted FRA's, but it can only be understood and risk managed as a change of modelling paradigm. And this model change broke down the hedging strategies of swaps (liquid and not liquid), the pricing of all swaps that were not liquid, and the foundations of the construction of the discounting curve for derivatives of any asset classes. All these hinged in fact on the above model assumptions. Thus, even what happens in liquid markets is strongly related to model risk, at least at a portfolio level. And, even worse: with the subprime crisis, the interbank market became suddenly illiquid... We do not really have "price uncertainty" in a liquid market, but we can have model risk and provisions against model risk, like reserves or position limits, can be useful also for liquid markets.

Regulators on Model Risk after the Crisis

Regulators have issued many documents and spent many words after the crisis. Some of them are relevant to model risk management and model validation. The approach they take after the crisis is very different from the schematic approaches suggested before the crisis. Many suggestions are reasonable and should be considered by modelers. How long before regulators (and we all) forget them again?

The most important document issued by regulators after the subprime crisis with a strong relevance for model risk is the Basel Committee's *"Supervisory guidance for assessing banks' financial instrument fair value practices"*, issued on 6 February 2009 (by far the most quoted doc at conferences on model risk and validation...). Below we follow this as a summary of up-to-date regulator views on modelling.

WHAT REGULATORS SAY AFTER CRISIS: BASEL in 2009

We recall below the main points contained in this document, and comment on them. There are three groups of recommendations (my classification).

The first group regards more the governance of the bank, and the organization of the risk management and model validation units. They regard the role of management in the model validation process, the use that must be made of its conclusions, the general principles that should be followed in the process.

The second group regards the relation between the model, the market and the specific product under analysis.

The third group of recommendations regard instead the operative steps that cannot be missed in a model validation process. Surprisingly enough, they are reasonable and fully in line with the above analysis (quants were involved in the writing of this document).

The idea should be, following Pillar II, that banks that stray too much from these principles will be required to keep more capital against model risk.

Basel new principles: general recommendations

1. *"Management must understand the basis of any valuation techniques used by outside parties... to determine the appropriateness of the techniques used, the underlying assumptions". "Reporting to senior management and to the board should be on a regular basis in an appropriately aggregated and understandable form"*
2. *"Bank valuation methodologies are expected to not place undue reliance on a single information source (eg external ratings) especially when valuing complex or illiquid products"*
3. *"Articulating the bank's tolerance for exposures subject to valuation uncertainty" "Analysis should be commensurate to the importance of the specific exposure for the overall solvency of the institution"*
4. *"Fair value measurements may involve a significant amount of judgment"*

Basel new principles: general recommendations

Points 1 and 2 stress that management has to be held responsible for the consequences of the valuation techniques used. This should mean that the discharge of responsibilities onto models and modellers that many managers did in the crisis should not be acceptable anymore... For this to make sense quants, modelers and risk managers should communicate regularly the choices made in valuation, and communication need to be understandable by the senior management. Seriously difficult... but needs to be done to avoid separation between those that understand models and those that take decisions based on them. Managers are responsible also when third party valuation is used. And the latter should not be taken at face value. This seems to recognize that excessive reliance on ratings was at the core of the crisis.

Basel new principles: general recommendations

Points 3 is common sense often lost in model validation. Model risk is not linear with respect to exposure. Mismodelling for small exposures is unavoidable (less resources, and banks learn from trial and error). Model sloppiness on large exposures is unacceptable, it can lead to dramatic reputational consequences, or even to default, in the worst case it can trigger a systemic crisis. This is the meaning of setting model reserves or model limits: not to allow exposures to grow too much when there is valuation uncertainty.

Point 4 requires no comment, but it is important to see it stressed out in an official document.

Similar points have been made at an earlier stage by FSA (March 2008 Report) and Financial Stability Forum (April 2008).

Basel new principles: the model, the market and the product

1. *"Uncertainty is specific to the instrument and to the point in time the valuation is effected"*
2. *"A bank is expected to consider all relevant market information likely to have a material effect on an instrument's fair value when selecting the appropriate inputs".*
3. *"However, observable inputs or transactions may not be relevant. In such cases, the observable data should be considered, but may not be determinative"*

Basel new principles: the model, the market and the product

These point regard again the relation with the market. All relevant market information must be used, but only the relevant one. First, there are some cases where market information can be of guidance but it is not binding (illiquid, biased, limited...) and then there are other cases where market information is simply not relevant for the instrument we are evaluating (we will see a case in considering gap risk and spread options). Including irrelevant market information only creates false confidence and hides valuation uncertainty.

Basel new principles: operative recommendations

1. *"Validation includes evaluations of the model's theoretical soundness and mathematical integrity and the appropriateness of model assumptions, including consistency with market"*
2. *"Bank processes should emphasise the importance of assessing fair value using a diversity of approaches and having in place a range of mechanisms to cross-check valuations"*
3. *"A bank is expected to test and review the performance of its valuation models under possible stress conditions, so that it understands the limitations of the models" "assess the impact of variations in model parameters on fair value, including under stress conditions"*
4. *"Policies should also identify specific triggers (eg indications of deterioration in model performance or quality) that will cause the review"*

Basel new principles: operative recommendations

Point 1 contains three steps in model validation: checking the mathematical computations, evaluate the soundness and appropriateness of its assumptions, check if the model is consistent with market prices. We have already seen them, and we will see them again.

Point 2 advocates a multiplicity of models and the need for model comparison. We will see this in detail for gap risk in credit, and will hint at that in equity and bermudan derivatives. No more emphasis on approximated models to be taken as the unique reference. This only fueled regulatory arbitrage.

Point 3 requires stress testing of the models to understand their limitations and stress testing with models to undestand the risks of a payoff. According to point 4, then, there should be quantitative triggers that allow us to say when the model reliability is deteriorating. We will see this for the rate market in 2007, for correlation models, and in the analysis of approximations. Stress-testing is the focus of many reg's docs in these days. We will discuss later what they mean by that.

Practical steps in Model Risk Management

We can make a summary of the points that emerged so far for a sound model validation.

Model Verification

1. Mathematics
2. Implementation
3. Numerics (to be tested under normal conditions)
4. Correct application to Payoff

Model Validation

1. Calibration
2. Reasonableness (assess that assumptions are not obviously unrealistic on relevant points).
3. Market Intelligence (to understand the model used by counterparties)
4. Reality check

* Detect the other models with equally sound assumptions and **compare the outputs**

- **Accept/reject the model** / **make improvements** / **Set model reserves and limits**

Practical steps in Model Risk Management

5 Once model has been chosen, apply **stress-testing**. This includes:

- verifying if the model can be used to express stress-conditions (**stressability**)

- subject the model to stress conditions to limit and monitor its application (**stress testing assumptions**: for **market conditions** and **payoff features**)

- detect what could impair the precision of analytics and numerics and test under these conditions (in particular **stress-testing approximations**).

* Analyze **how the current market consensus (or lack of it) could change in the future** and how this affects the model use

- **Set model reserves and limits/ Set monitor Triggers**

Practical examples in Model Risk Management

Other issues may complement the analysis:

- **Hedging with the model**: 1) assess if the model-implied hedging can give some indications on model validity (P&L analysis) 2) the real hedging strategy can be different from the model implied one and may need a separate validation (Hedge analysis).
- **Extrapolations and Interpolations** can be used as a supplement to calibration. This is an additional model risk. The use of extrapolation can be avoided or minimized, interpolation can be checked out.
- Some models are used to perform **model or statistical arbitrage**. The meaning and the practical implications of this sort of application must be understood and assessed.

Comparing Models: an analysis on Credit Models and Gap Risk

We analyze how we can use different models in order to study the model risk that affects a payoff. We have chosen a payoff which appears simple, and we limit ourselves to two modelling approaches which are both popular and largely used in the market. Yet there is heavy and relevant model uncertainty.

The two popular models represent two opposite assumptions about the market behaviour with respect to the payoff we consider. Understanding this requires first an analysis of the core risk in the payoff. Then we see that there is no liquid available information on market consensus, nor there is unambiguous historical evidence, so that valuation unavoidably depends on judgement. We analyse the models representing the two limit cases for the risk of the product, and construct a family of intermediate models. This works as a stress test on the value of the payoff and allows to understand better the different assumptions that different models make about it. The analysis will give useful elements to make a choice on the model we want to use and then to quantify the residual model risk, to set a knowledgeable model reserve or a model limit line.

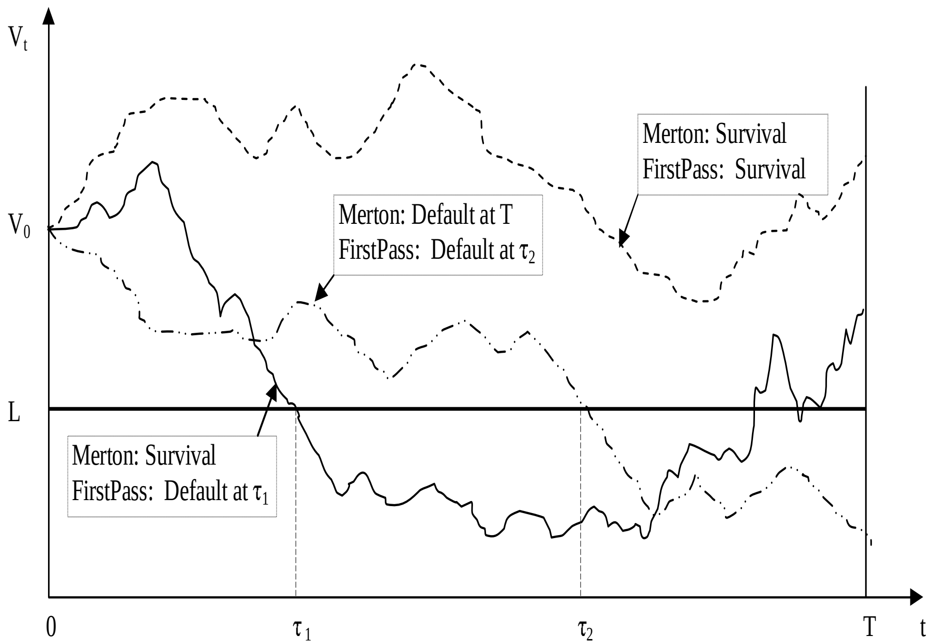
Credit Derivatives Modelling. Approaches

Let us recall the main two families of models for credit risk.

- **Structural Models** (Balance sheet quantities, such as firm value and debt, are the fundamental modelling quantities. The models represent default as an economic event, the causes are considered explicitly)
- **Reduced-Form or Intensity Models** (Default intensities are the fundamental modelling quantities. The models represent default as an exogenous, unpredictable event, the causes are not considered explicitly)

Black and Cox approach: First Passage Models

Default happens when the **value of the assets** V_t first hits from above a **default barrier** H, associated to firm's debt



Structural Models: Black Cox Model

If V_t goes below the barrier H there is default. V is again a standard Geometric Brownian Motion

$$dV_t = (r - q)\, V_t dt + \sigma V_t dW$$

Default happens at

$$\tau = \inf\{t | V_t \leq H_t\},$$

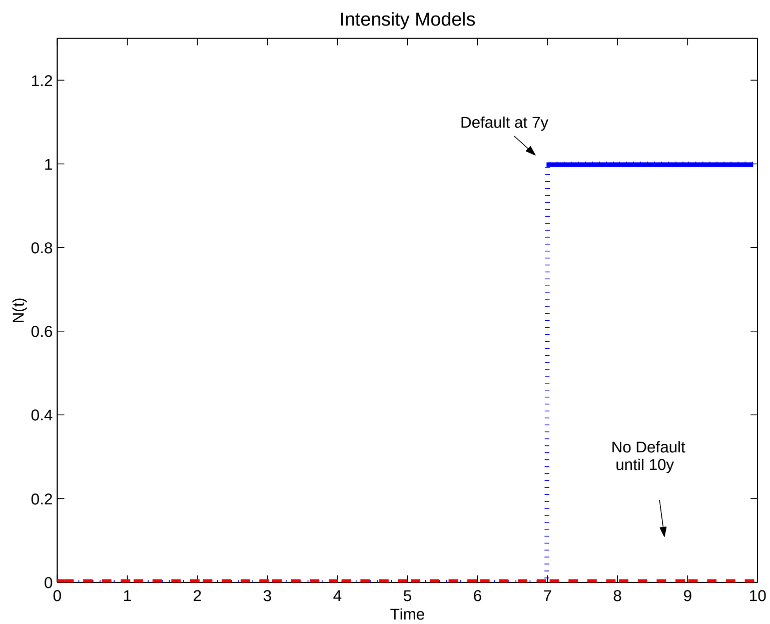
the **first passage** of the firm value below the barrier. Assuming a fixed barrier $H_t = L$, the situation is similar to a **barrier option** (down and out). It is fundamental to compute the probability of not touching the barrier, in credit this is survival probability:

$$Q\{\tau > T\} = N\left(\frac{\ln\frac{V_0}{L} + \left(r - q - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right) - \left(\frac{V_0}{L}\right)^{-\frac{2\left(r-q-\frac{1}{2}\sigma^2\right)}{\sigma^2}} N\left(\frac{-\ln\frac{V_0}{L} + \left(r - q - \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}\right)$$

Brigo and Morini (2006) give analogous closed-form formulas also for time-dependent parameters and barrier, allowing calibration to CDS term structure.

Intensity Approach: Poisson Processes

Intensity Models (Default intensities are the fundamental modelling quantities. The models represent default as an exogenous, unpredictable event, the causes are not considered explicitly). Default of an individual name happens when a **jump process** N_t jumps for the first time



Intensity Models

Intensity Modelling

The fundamental assumptions of intensity models are

A the probability of 1 event in Δt is λ proportional to Δt.

B Δt is so small that there cannot be 2 events in Δt.

C events in disjoint intervals are independent.

Intensity Modelling

In single name modelling we will take τ^1, the first jump of the poisson process, as default time τ. Based on the above assumptions, we can give the distribution of default time:

$$\begin{aligned} Pr(\tau > T) &= e^{-\lambda T}, \\ F_\tau(T) &= Pr(\tau \le T) = 1 - e^{-\lambda T}, \end{aligned}$$

which a **(negative) exponential distribution**, with parameter λ, so

$$\boxed{\tau = F_\tau^{-1}(U) = -\tfrac{1}{\lambda}\ln(1-U).}$$

If we set $\tau\lambda = \varepsilon$, we see ε is a unit exponential rv, since

$$F_\varepsilon(z) = \Pr(\varepsilon \le z) = F_\tau\left(\frac{z}{\lambda}\right) = 1 - e^{-z}.$$

So I can also write

$$\tau = \frac{\varepsilon}{\lambda}$$

Time-dependent credit spreads

With time inhomogeneous Poisson process we define

$$\Lambda(T) = \int_0^T \lambda(s)\, ds$$

For τ, we have that we can write it as

$$\tau = \Lambda^{-1}(\varepsilon) = \Lambda^{-1}(-\ln(1-U))$$

with ε a unit exponential rv and U a uniform rv. The generalized inverse, that we use in simulation also with stochastic intensity (on every single path), is

$$\tau = \inf\left\{t : \int_0^t \lambda(s)\, ds \geq \varepsilon\right\}.$$

Credit Derivatives Modelling. Approaches

In credit modelling it can be meaningful to separate default free information from information containing the default event:

$$\mathcal{F}_t = \mathcal{H}_t \vee \mathcal{F}_t^{\tau}$$

- $\mathcal{F}_t =$ all available information up to t
- $\mathcal{F}_t^{\tau} = \sigma(\{\tau < u\}, u \leq t) =$ information up to t on the default event: if it has already happened or not, and in the former case the exact time τ of default
- $\mathcal{H}_t =$information up to t on economic quantities which affect default probability, but no specific information on happening of default

In standard **structural models** the distinction is meaningless, we would have: $\mathcal{H}_t = \mathcal{F}_t$ since same economic quantities inform both on default probabilities and default time
Instead in standard **reduced form models** the distinction is fundamental: $\mathcal{H}_t \subset \mathcal{F}_t$ since default is exogenous to economic quantities, which only affect default probability.
In **market models** too $\mathcal{H}_t \subset \mathcal{F}_t$ appears the correct choice.

Pricing a Note with a Trigger

Consider a note issued by a bank, with a notional$= 1$, to sell leveraged protection on a reference entity. The client pays 1, to be given back at maturity T if the reference entity survives. The bank sells protection for a leveraged notional Lev and pay to the client a quantity

$$Lev \times S_T(0),$$

where $S_T(0)$ is the spread of the reference entity, minus... what? In case of a default, the maximum the client can lose is 1. But the loss in case of default may be higher. The gap

$$Lev \times Lgd - 1$$

is covered by the bank, that therefore suffers a *gap risk*, expressed by an expected loss

$$\mathbb{E}\left[1_{\{\tau<T\}} P(0,\tau) \times (Lev \times Lgd - 1)^{+}\right]. \tag{5}$$

where $P(0,\tau)$ is the discount factor, and this reduces the spread paid to the client.There can be different ways to mitigate the gap risk.

Pricing a Note with a Trigger

One popular way is to set a *trigger*, namely**:** *the note is interrupted and settled at market value before maturity when*

Event A there is default event on X (we call the default time τ)

Event B the value at t of $S_T^X(t)$, the spread of the CDS with maturity T, touches a level $trigger$ (we call the touch time t^{touch})

Thus the early termination date is at

$$T^{term} = \min\left(\tau, t^{touch}\right)$$

The issue here is to quantify Gap Risk for the bank now.

Pricing a Note with a Trigger

The bank has a gap loss when we jointly have

$$T^{term} < T \ \ \textbf{AND} \ \ Loss(T^{term}) > 1$$

Thus the NPV of the possible loss associated to Gap Risk is

$$\mathbb{E}_t \left[1_{\{T^{Term} < T\}} P\left(t, T^{Term}\right) \times \left(Lev \times CDS_T\left(T^{Term}\right) - 1 \right)^+ \right]$$

where $CDS_T\left(T^{Term}\right)$ is the market value at T^{Term} of the CDS with maturity T, with a fixed spread $S_T(0)$. The financial rationale of the note, for the issuing bank, is that in case of worsening credit conditions, the presence of a trigger should interrupt the note **before** default time (at $t^{touch} < \tau$), with a loss $Lev \times CDS_T\left(t^{touch}\right) < Lev \times CDS_T(\tau) = Lev \times Lgd$.

Pricing a Note with a Trigger

The bank can set the trigger at a level $trigger^*$ such that, when the note is unwound at $trigger^*$ we have a loss equal to 1, which is the notional of the leveraged note. In this way the loss will be entirely covered by the client and the bank has no gap loss. How can we set the level $trigger^*$? There is a relation between the value of a CDS at t and the spread $S_T(t)$. For example, using a simple approximation,

$$\Pr_t(\tau > s) = \exp\left(-\frac{S_T(t)}{Lgd} \times (s-t)\right),$$

and with this we can find the CDS value as a function of the spread,

$$CDS_T(t) = f(S_T(t)).$$

and compute

$$trigger^* = f^{-1}\left(\frac{1}{Lev}\right).$$

There are two fundamental families of credit models that banks use for simple credit derivatives like the one above. They are structural models, the first approach to credit risk, and reduced-form models where the crucial variable is the default intensity.

First test: basic calibration

A first discrimination can be based on the capability to calibrate the relevant financial information. Surely this product depends on default probability, thus the model must be able to calibrate the CDS term structure.

Intensity models, even in the standard form with deterministic, time-dependent intensity, have a good capability to calibrate CDS term structures.

Standard structural models, however, cannot do this and can be ruled out from our set of eligible models.

But there are modern structural models, mentioned above, that have time-dependent parameters and analytical solutions so they allow to calibrate CDS as well as intensity models.

Second test: minimum realism

Secondly, there are some basic features of realism that one would like a model to include when applied to the above derivative. The unwinding of the note above happens when the credit spread *touches* a trigger. Spreads should have a stochastic dynamics for the derivative to make sense. For structural models this is not a problem. The value of the firm A_t evolves stochastically as

$$dA_t = (r - q)\, A_t dt + \sigma A_t dW.$$

The credit spreads at a future date $s > 0$ are functions of the survival probability at s,

$$\Pr_s(\tau > T) = N\left(\frac{\ln\frac{A_s}{L} + \left(r - \frac{\sigma^2}{2}\right)(T-s)}{\sigma\sqrt{(T-s)}}\right) - \left(\frac{A_s}{L}\right)^{-\frac{2\left(r-\frac{1}{2}\sigma^2\right)}{\sigma^2}} N\left(\frac{-\ln\frac{A_s}{L} + \left(r - \frac{\sigma^2}{2}\right)(T-s)}{\sigma\sqrt{(T-s)}}\right).$$

This probability at future time s depends on A_s which is not known at zero but will depend on the stochastic dynamics given above. Thus the entire quantity $\Pr_s(\tau > T)$ varies stochastically with A_s, and credit spreads vary accordingly.

Second test: minimum realism

What about intensity models? In standard intensity models the intensity $\lambda(t)$ is deterministic. The survival probability

$$Q(\tau > T | F_t) = e^{-\int_t^T \lambda(s)ds}$$

is also deterministic, and so is the spread. Thus it is possible to know at time 0 if the trigger will be touched or not, and in case it will be touched we know exactly when this is going to happen. This is an element of lack of realism that strongly affects our valuation, since the moment when we touch the spread is crucial to assess the Gap risk. Clearly one needs a stochastic intensity model to move to a more realistic behaviour allowing for spread volatility. Intensity dynamics is often modelled analogously to the instantaneous spot rate in early interest rate models.

Stochastic Credit spreads in intensity models

First of all we have to rule out lack of credit spread volatility intensity models. Modelling the volatility of credit spreads amounts to assume that default intensities are stochastic. Poisson processes with stochastic intensity are known as Cox Processes

$\lambda \longrightarrow$	$\lambda(t) \longrightarrow$	$\lambda(t,\omega)$
Poisson Process	Inhomogeneous Poisson Process	Cox Process

In stochastic intensity models

$$\Pr(\tau \leq T) = \mathbb{E}\left[e^{-\int_0^T \lambda(s)ds}\right]$$

Stochastic Credit spread: a practical example

In a practical model, one would like $\mathbb{E}_t\left[e^{-\int_t^T \lambda(s)ds}\right]$ to have an analytic solution. Notice it has got the same form as the bond price in short interest rate models. Thus spread dynamics is often modelled analogously to the instantaneous spot rate in early interest rate models. If we choose the CIR (Cox, Ingersoll, Ross) dynamics

$$d\lambda\left(t\right) = k[\theta - \lambda\left(t\right)]dt + \sigma\sqrt{\lambda\left(t\right)}dW\left(t\right),$$

one can guarantee positivity of the spreads through $\sigma^2 < 2k\theta$, and the survival probability can be explicitly computed as

$$\mathbb{E}_t\left[\exp\left(-\int_t^T \lambda\left(u\right)du\right)\right] = A(t,T)e^{-B(t,T)\lambda(t)},$$

Stochastic Credit Spreads in Practice: Intensity Models

$$\mathbb{E}_t\left[\exp\left(-\int_t^T \lambda(u)\,du\right)\right] = A(t,T)e^{-B(t,T)\lambda(t)},$$

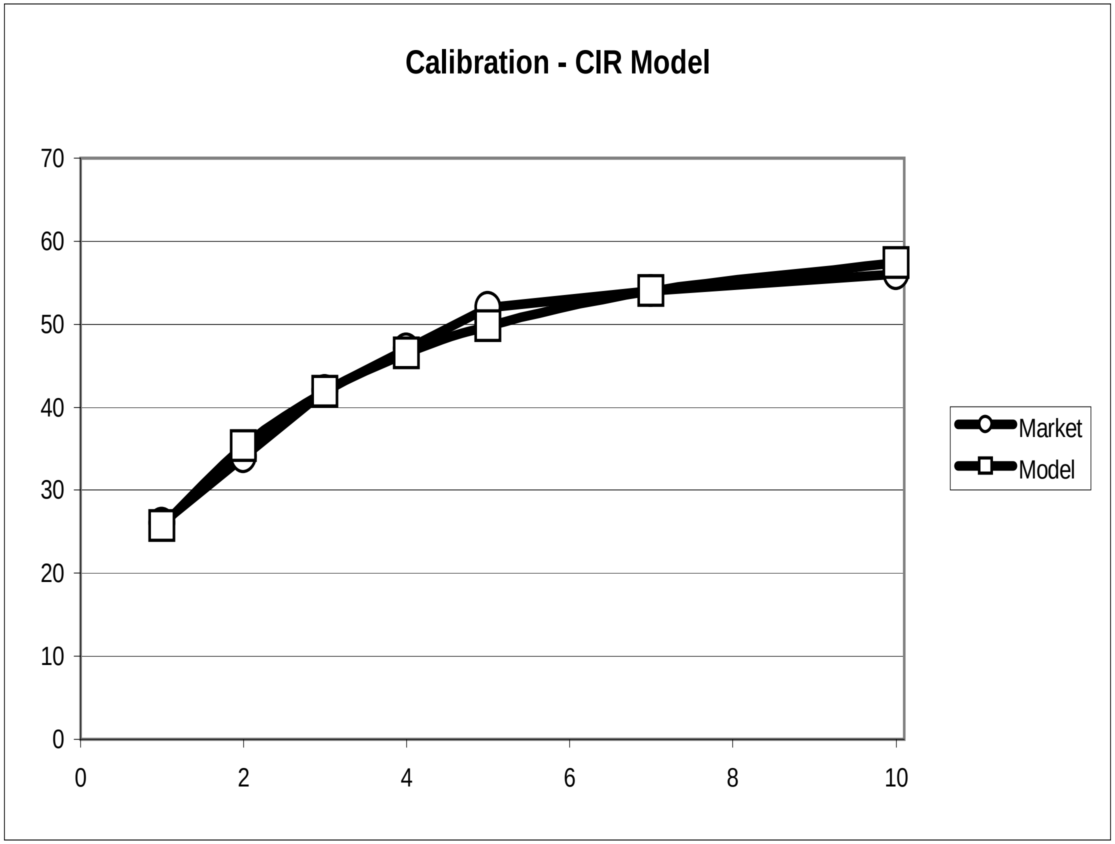
$$\begin{aligned}
A(t,T) &= \left[\frac{2h\exp\left[(k+h)(T-t)/2\right]}{2h+(k+h)\left[\exp\left((T-t)h\right)-1\right]}\right]^{2k\theta/\sigma^2} \\
B(t,T) &= \frac{2\left[\exp\left((T-t)h\right)-1\right]}{2h+(k+h)\left[\exp\left((T-t)h\right)-1\right]} \\
h &= \sqrt{k^2+2\sigma^2}
\end{aligned} \tag{6}$$

This makes calibration to CDS very easy. Since most of the times there are no liquid options on CDS, CDS calibration is the only calibration available. The model has four parameters,

$$\lambda(0)\,,\theta,k,\sigma$$

so a standard CDS term structure can be sufficient for stable calibration.

Stochastic Credit Spreads in Practice: Intensity Models



Next Steps

Are we sure this is sufficient? Are there other important elements in the spread behaviour that we have to assess? Is the calibration process concluded with CDS? Are now these two families of models perfectly equivalent in assessing our gap risk?

In order to answer these questions we need to perform a deeper analysis of the payoff.

Pricing a Note with a Trigger

The gap risk in the product becomes a loss when the reference entity defaults. This is avoided if, before the default, the spread increases so much that the trigger is touched. Thus the value of the Gap Risk crucially depends *on the behaviour of spreads in the time immediately preceding default*: will default be preceded by a significant rally in spread, or we will have a more abrupt leap to default?

In the first case, before default the note will be terminated with a cost that should be covered by the note's notional. The only possible loss for the issuer arises when the growth of the spread is so fast that when the actual unwinding takes place the spread S_T^X is much higher than Trigger. But if the unwinding is timely the trigger condition can eliminate gap risk.

In the second case, when default is not preceded by a spread rally, the trigger will have no effect. The bank will surely suffer a loss, equal to

$$Lev \times CDS_T(\tau) - 1 = Lev \times (1 - Rec) - 1.$$

Structural Models and Gap Risk

In Structural models, if the value of the firm A_t approaches the default barrier L, the survival probability

$$\Pr\left(\tau > T | A_t\right) = N\left(\frac{\ln\frac{A_t}{L} + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}\right) - \left(\frac{A_t}{L}\right)^{-\frac{2\left(r-\frac{1}{2}\sigma^2\right)}{\sigma^2}} N\left(\frac{-\ln\frac{A_t}{L} + \left(r - \frac{\sigma^2}{2}\right)(T-t)}{\sigma\sqrt{(T-t)}}\right)$$

falls, since default is given by the value of the firm touching the barrier and the lower the value of the firm the more likely this becomes,

$$A_t \downarrow \implies \Pr\left(\tau > T | A_t\right) \downarrow .$$

The survival probability converges to 0 when $A_t \approx L$, driving to extremely high values the CDS spread

$$S_{T_b} = \text{LGD} \sum_{i=a+1}^{b} Q\left(T_{i-1} < \tau \leq T_i\right) P(0,T_i) / \sum_{i=a+1}^{b} P(0,T_i)\alpha_i Q\left(\tau > T_i\right)$$

Structural Models and Gap Risk

Thus in Structural Models default will always be preceded by a spread really. If the notes involves a trigger, at any level, it will be touched before default. If liquidation can be performed timely, the note will be liquidated before suffering the default loss, and if the trigger as been set lower than or equal to $trigger^*$ there is no Gap Risk. Even taking into account a non-continuous but daily monitoring of the spread, we remain around

$$GapRisk \approx 0.$$

The 'spread to infinity' prediction of these structural models implies an absolutely extreme, and very aggressive, evaluation of the Gap risk.

Intensity Models and Gap Risk

If $\lambda(t)$ is deterministic we can predict all future spread levels. If the trigger is set at threshold higher than any predictable spread, no doubt the trigger will never be touched. We are at the opposite end compared to structural models, the gap risk is maximum, the same we would have with no triggers,

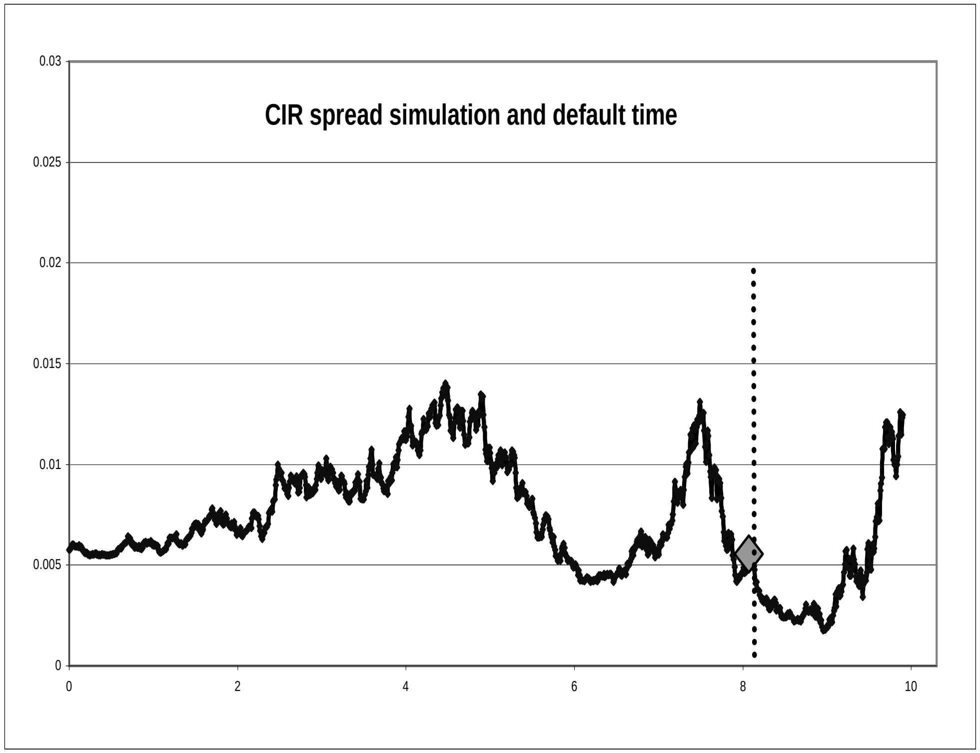
$$GapRisk = MaxGapRisk = Discount \times \Pr(\tau < T_M) \times (Lev \times (1 - \operatorname{Re} c) - 1) \tag{7}$$

We have introduced stochastic intensity via a process

$$d\lambda(t) = k[\theta - \lambda(t)]dt + \sigma\sqrt{\lambda(t)}dW(t).$$

that gives spreads a non-trivial dynamics. What is the behaviour of the spreads in proximity of default implied by such a model? This is what really matter for the gap risk. Let us have a look at one typical path for the spread, with the default time indicated by the dotted vertical line.

Intensity Models and Gap Risk



Intensity Models and Gap Risk

We see that the default time does not appear to be associated to an increase in spreads. In this chart default happens even after a spread decline. It is the typical *jump or leap to default*. One may wonder if the above montecarlo scenario is not an outlier... unfortunately is not. The high likelihood of jumps to default in an intensity model like the one described above is associated to well documented features of intensity models.

The low 'spread-default correlation' in intensity models with diffusive stochastic intensity had already been noticed, in other contexts (Roncalli et al., 2001). We can understand it also analytically. In intensity models with diffusive intensity the Gap risk will tend to its maximum values, the one that the bank would bear if there was no trigger:

$$GapRisk \approx MaxGapRisk = Discount \times \Pr\left(\tau < T_M\right) \times \left(Lev \times \left(1 - \operatorname{Re} c\right) - 1\right).$$

Aren't we making a mistake opposite to structural models?

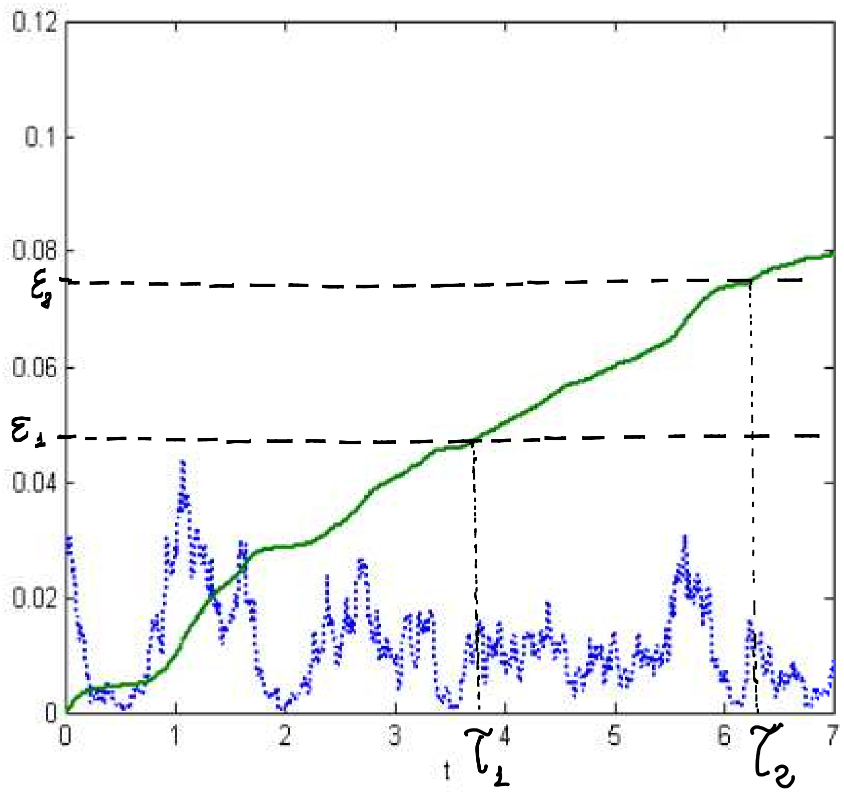


Figure 1: Diffusive intensity. The dotted line is a montecarlo path, the continuous line is its integral, the dashed lines are the default triggers.

Model Uncertainty: market information

Are there any liquid products in the market that can allow us to solve this indetermination, giving us an indication on the market assessment about this gap risk, allowing us to chose between a structural model or an intensity model? For some reference entities it is possible to find relatively liquid markets for CDS option. Options in the market are knock out, namely they expire anticipatedly if the company defaults. Thus their spread-dependent payoff $(S - K)^+$ clearly depends on the dynamic behaviour of spreads, but *conditionally on having no default during the life of the option*. Here we are interested to know, instead, the dynamic behaviour of the spreads conditional on *being in the very proximity of default*. The two derivatives speak of two different states of the world. Options provide useful information but do not give the crucial information on the bank's gap risk.

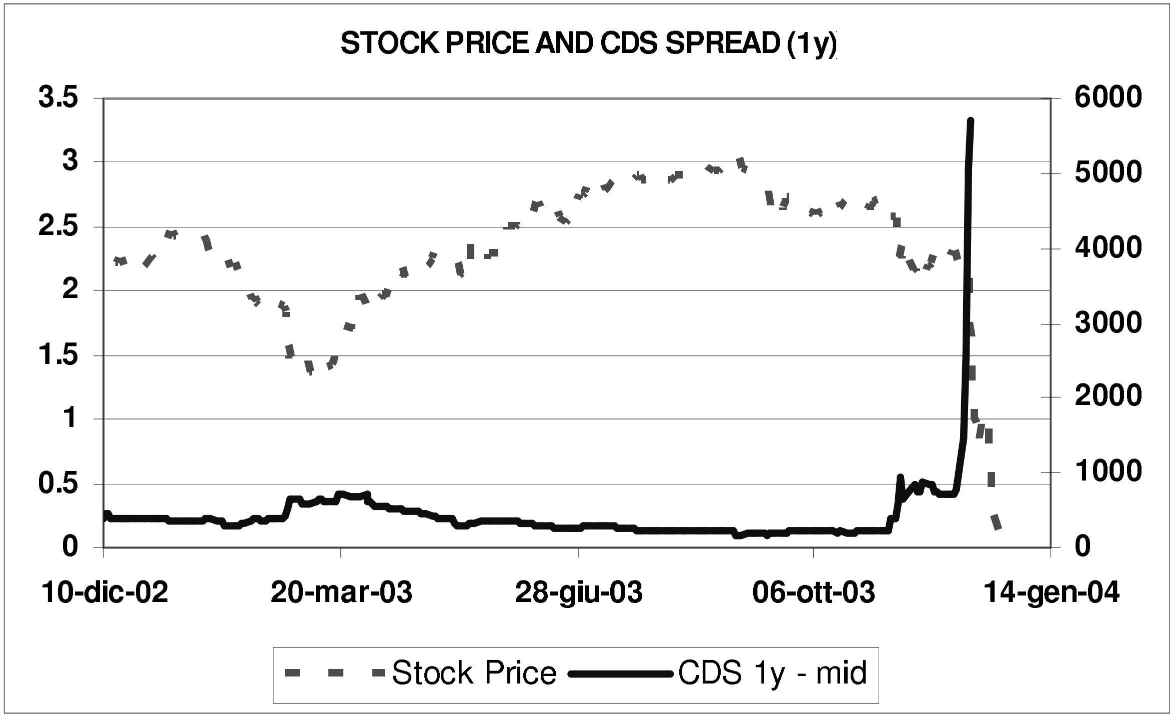
Model Uncertainty: incompleteness

The leveraged note, however a simple product, depends crucially on a risk factor, driving the relation between spreads and default time, which is not observable in any market because of the lack of a liquid derivative depending on it. Thus the market is *incomplete*. In an incomplete market, even if all market participants agree on the real world probabilities of an event (for example on the real world probability of being saved from gap losses by a spread rally that touches the trigger) they can disagree on the compensation they want for the risk of this event. A conservative bank will increase the risk neutral likelihood of a jump-to-default used in pricing. We can't say if the divergence is based on different expectations about the default behaviour, or different risk aversions.

Model Uncertainty: historical evidence

Do we have a clear historical evidence that can give us an indicator of what is the most likely possibility, a jump to default or a default preceded by a relatively smooth spread increase? The first caveat is that history of past defaults has little to do with future defaults. And also the historical evidence is mixed. A pure jump to default is not common, but Enron and Parmalat were almost like that. Argentina, instead, a sovereign default, was fully predictable, with spreads growing steadily until a climax at default. Lehman was mixed. In general, the likelihood of a sudden default is inversely proportional to the transparency of the balance sheet of the reference entity.

Parmalat before default



Parmalat stock price (in euros, left scale) and 1y-CDS spread (in bps., right scale).

Intermediate Models

Can we find a model where the behaviour is somewhat in-between the null Gap Risk of Structural Models and the maximum Gap Risk of intensity models? There are different possibilities. The first one we can think of is to make one step further in increasing the realism of structural models, and take Structural Models with a default barrier that can jump between different levels. An unpredictable barrier would create also in this models a possibility of a jump to default, and we could play with the parameters to control its probability. An alternative, that we will follow in more detail because it is somewhat more original and also more tractable, are intensity models where stochastic intensity *which is not anymore a diffusion but* can experience jumps.

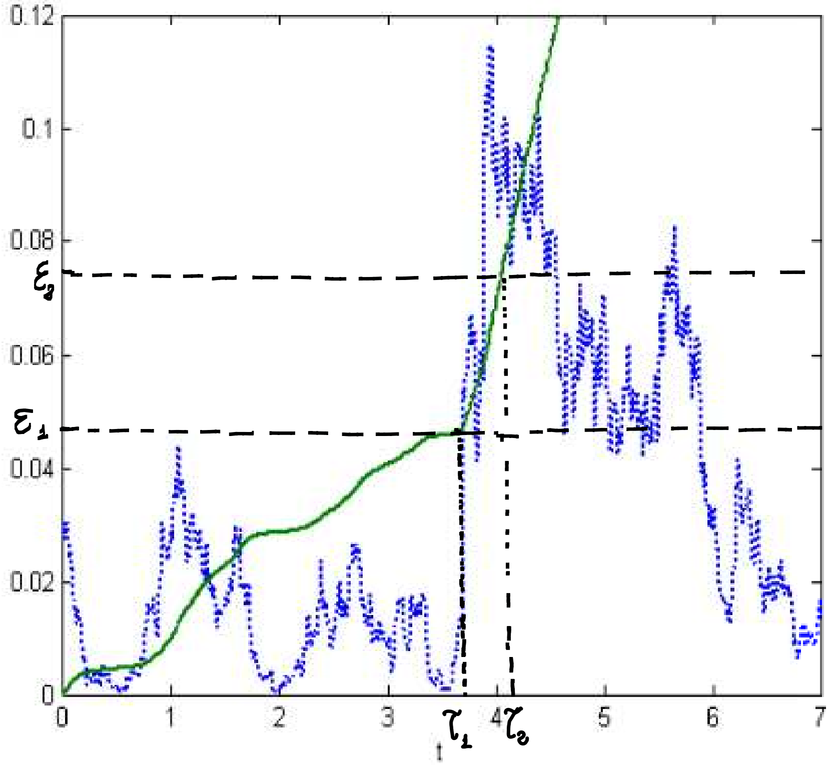


Figure 2: Jump-diffusion intensity. The dotted line is a montecarlo path, the continuous line is its integral, the dashed lines are the default triggers.

Jumps in the Spreads

We consider the Cir-exponential Jump model

$$d\lambda(t) = k[\theta - \lambda(t)]dt + \sigma\sqrt{\lambda(t)}dW(t) + dJ_t^{\alpha,\gamma}.$$

where

$$J_t^{\alpha,\gamma} = \sum_{i=1}^{N_t} Y_i$$

with N_t a Poisson process with jump-intensity α and $Y_i \sim \exp(1/\gamma)$, being γ the expected jump size.

Jumps in the Spreads

With this choice of jump distribution, survival probability can be computed analytically

$$\mathbb{E}_t\left[\exp\left(-\int_t^T \lambda\left(u\right)du\right)\right] = \bar{A}(t,T)e^{-\bar{B}(t,T)\lambda(t)},$$

where

$$h \neq k + 2\gamma :$$

$$\bar{A}(t,T) = A(t,T)\left[\frac{2h\exp\left[\left(k+h+2\gamma\right)\left(T-t\right)/2\right]}{2h+\left(k+h+2\gamma\right)\left[\exp\left(\left(T-t\right)h\right)-1\right]}\right]^{\frac{2\alpha\gamma}{\sigma^2-2k\gamma-2\gamma^2}}$$

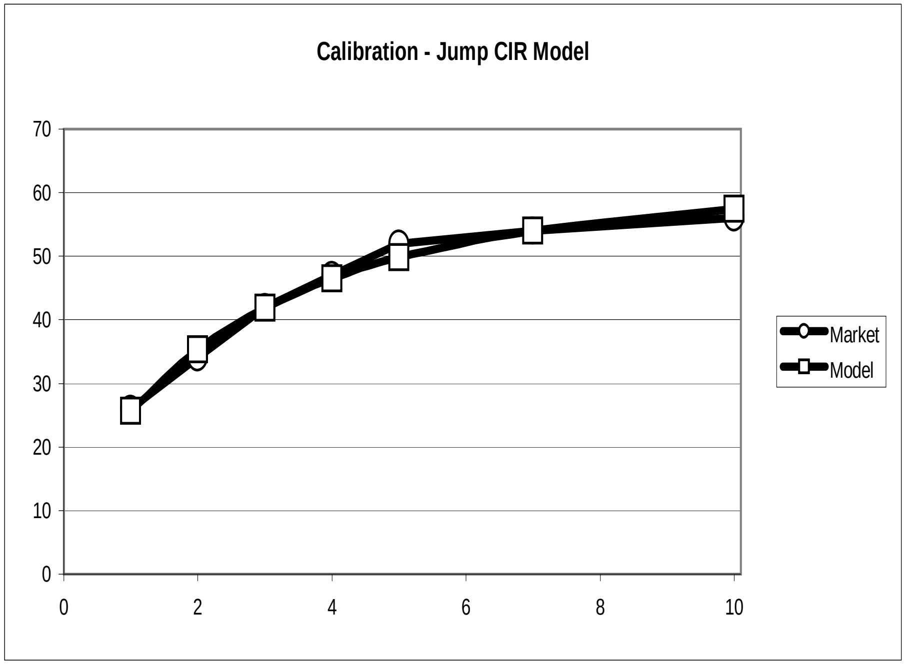
$$h = k + 2\gamma :$$

$$\bar{A}(t,T) = A(t,T)\exp\left[-2\alpha\gamma\left(\frac{T-t}{k+h+2\gamma}+\frac{\exp\left(\left(T-t\right)h\right)-1}{h\left(k+h+2\gamma\right)}\right)\right]$$

$$\bar{B}(t,T) = B(t,T)$$

Jumps in the Spreads

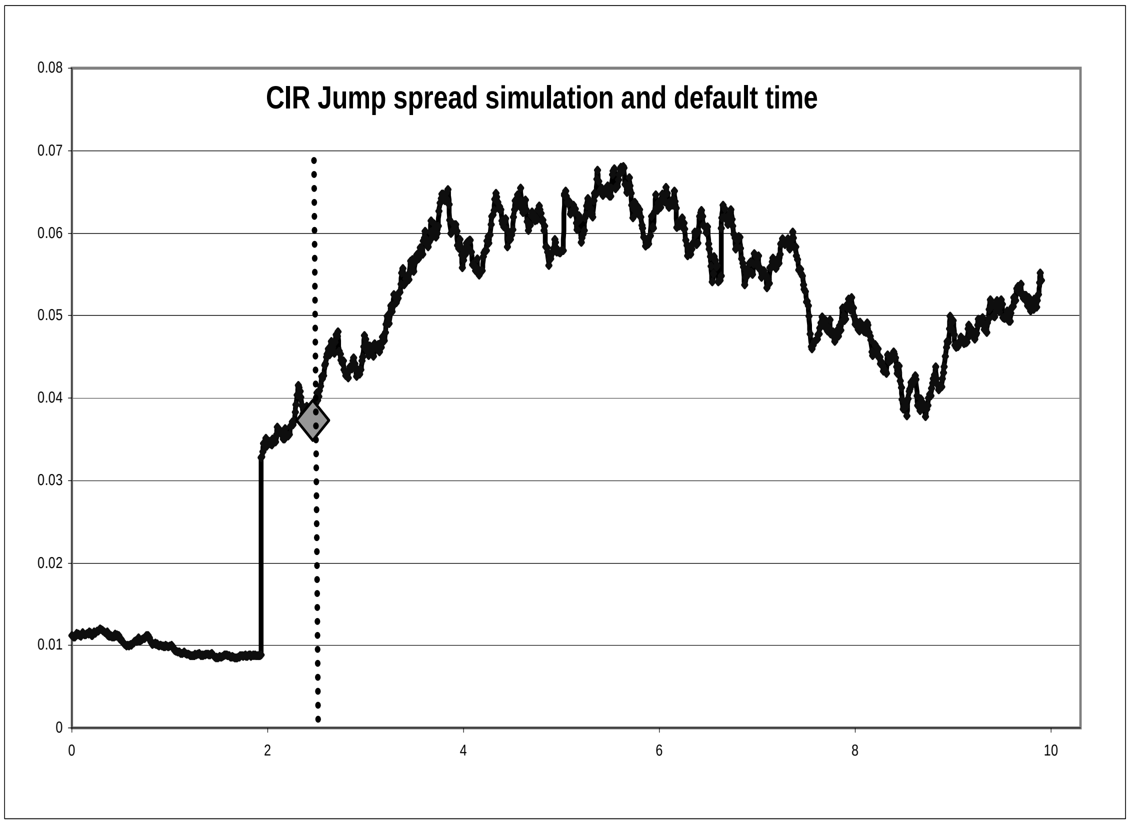
This makes calibration simple. For example, using the same CDS term structure used for the diffusive model, but imposing the presence of jumps with $\gamma = 100bps$, we get



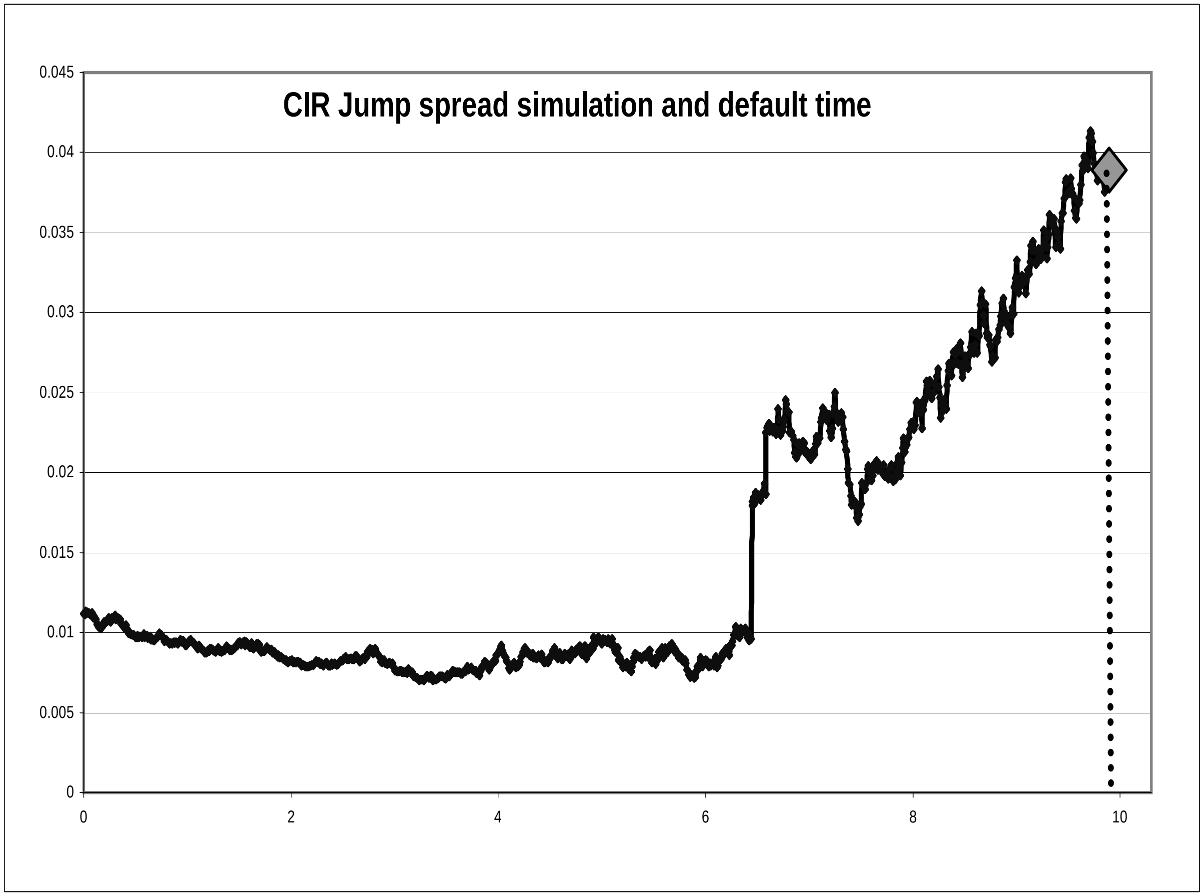
One can notice calibration does not change adding the jump parameters. This confirms that, without additional single name calibration products, both a diffusive and a jump model can be consistent with CDS (market incompleteness).

Jumps in the Spreads

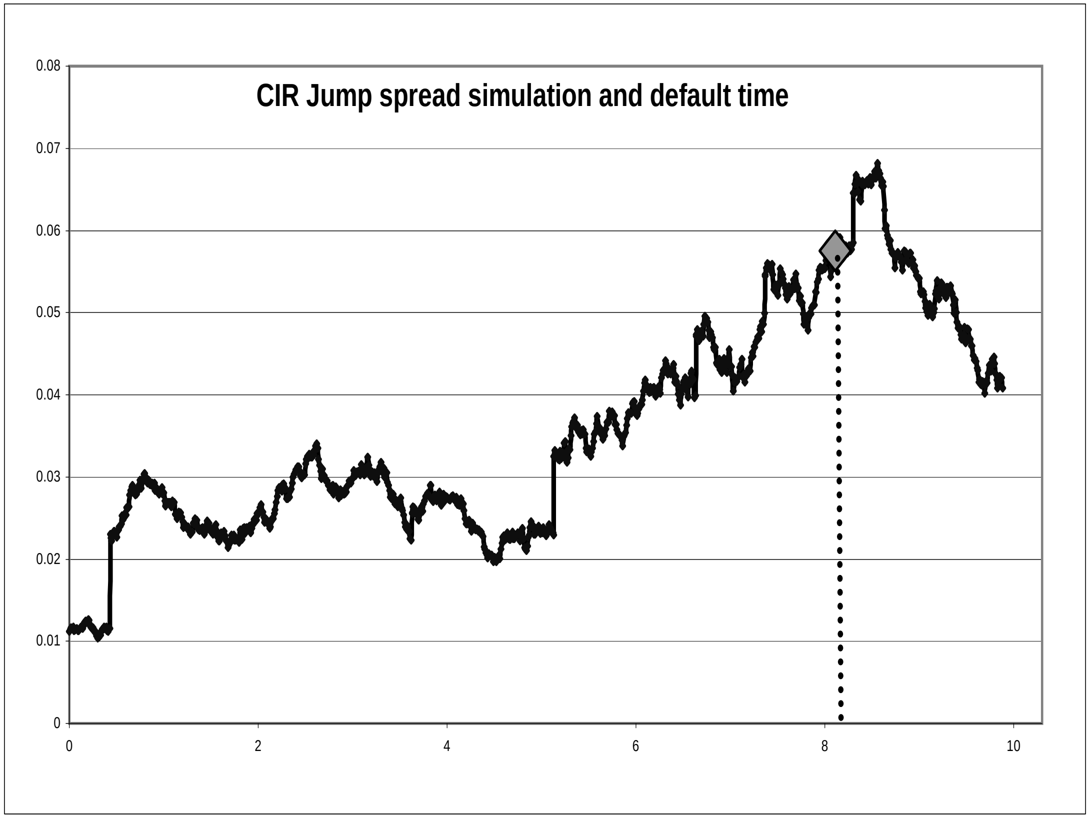
However, the behaviour of the model including spread jumps ($\gamma = 0.1$) is very different, in terms of relations between spread level and default. Now we have more correlation between default and very high spread.



Jumps in the Spreads



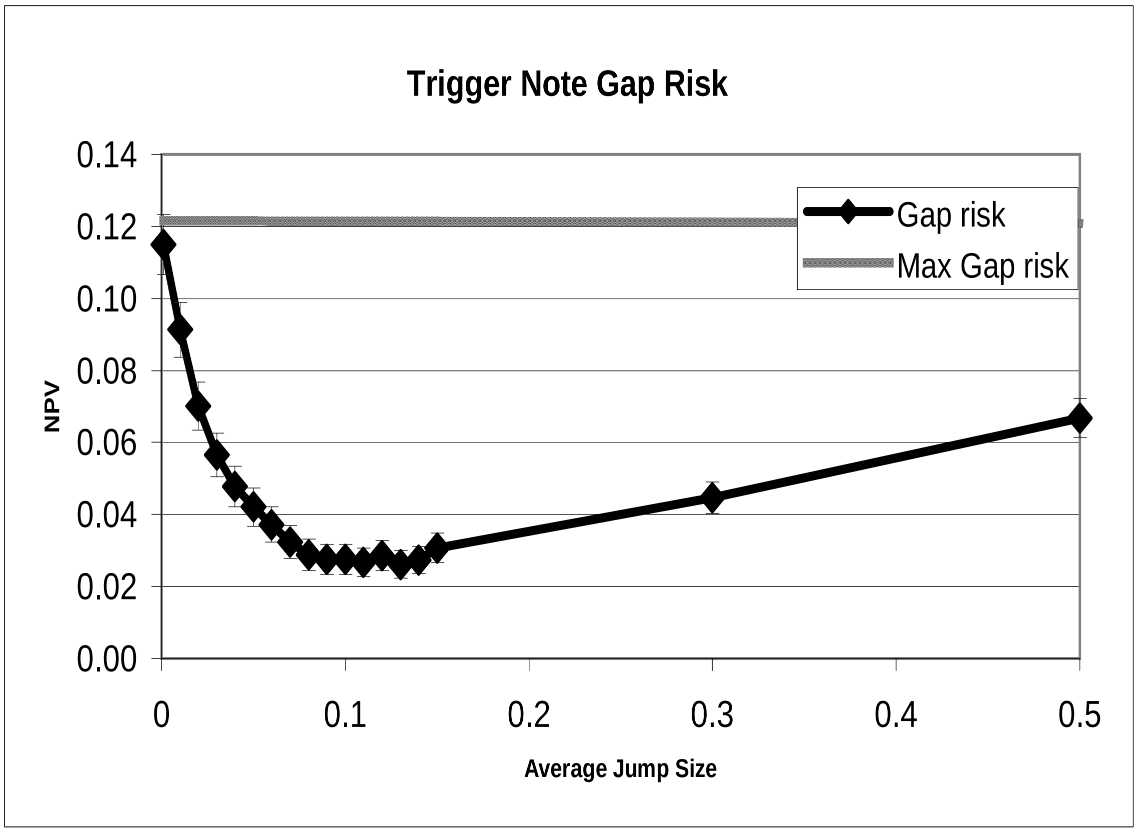
Jumps in the Spreads



And this reflects into the value of the gap risk...

Jumps in the Spreads

Take a note with $T_M = 5y$ and $trigger = 2.5 \times S_T^X(0)$, we have the following pattern for Gap Risk moving from no jumps to the possibility of very high jumps (see Facchinetti and Morini (2008))



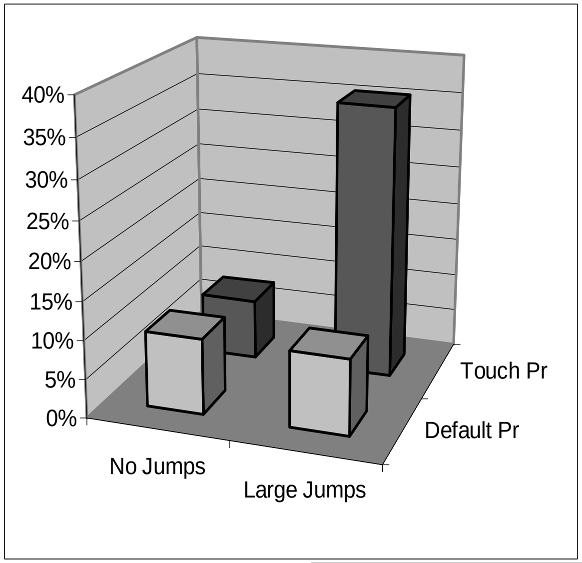
Pricing a Note with a Trigger

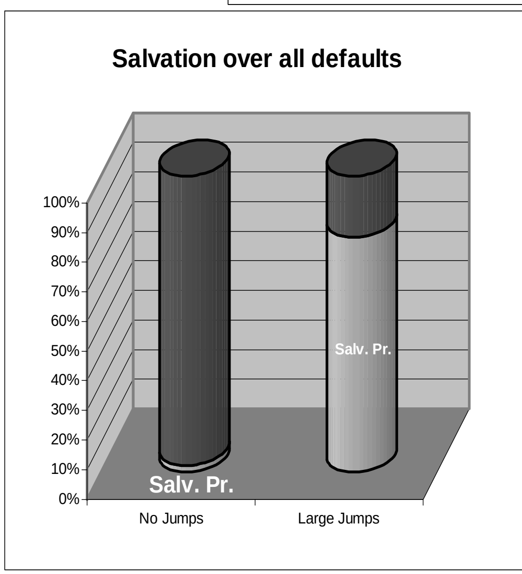
Notice that in all above cases the model has been calibrated to the same CDS data, and always implies (almost) the same default probability. Notice also that the model does not allow to send the gap risk to zero, the possibility of totally unexpected defaults (or so high and sudden spread rallies that lead in any case to losses for the bank) is always included.

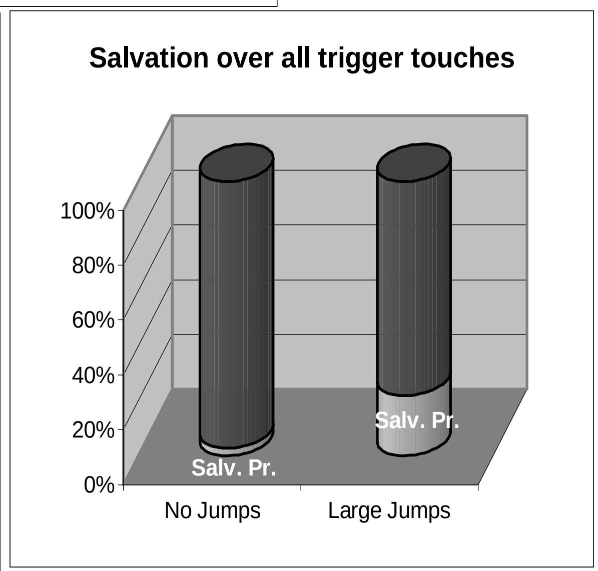
The reason for the differences in Gap Risk appears even clearer if we compare two extreme configuration in terms of the probability of touching the trigger before maturity (Touch Probability). It can be seen in the top chart of the figure below, considering no jumps vs large jumps$\approx$1000bps (is such a jump unrealistic? Consider that for us it has the meaning of the intensity/spread jump that can lead to default, so that it is not unrealistic).

Pricing a Note with a Trigger

In order to understand how differently the two extreme models represent the Gap Risk and the note financial meaning, it is interesting to see how many defaults that would have led to high gap risk for the bank were actually avoided by touching the trigger. This is the "Salvation Probability" of the chart "Salvation over all defaults" on the left of the figure below (Pre-Default-Touches over all Defaults). We also provide another relevant information. We decompose scenarios when the touch actually "saved" the note from default from scenarios when the touch stopped the note but than there was no default of the reference name before maturity (a very negative situation for the holder of the note). This is the "Salvation Probability" of the chart "Salvation over all trigger touches" on the right of the figure below (Pre-Default-Touches over all Touches). Notice that in the test reported in the figure we are using two models calibrated to the same CDS data, as indicated also by the fact that the default probabilities are the same.

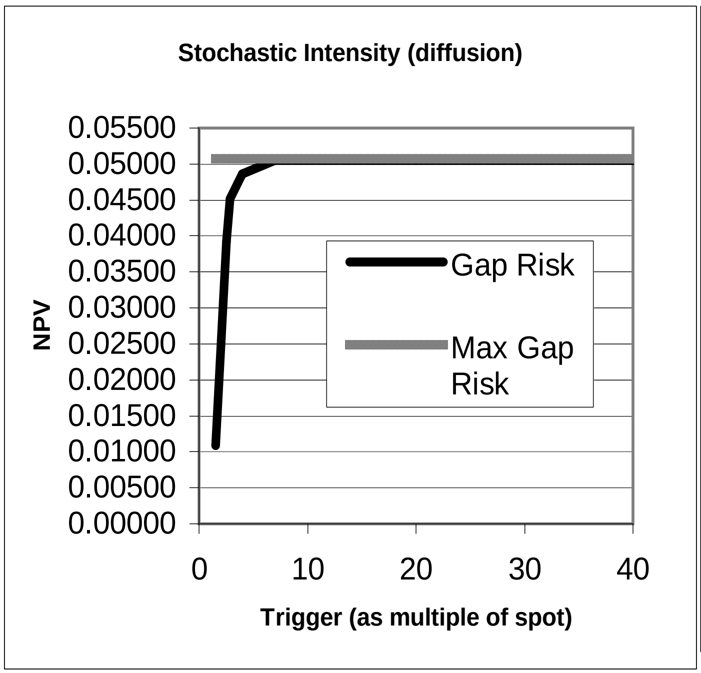

40%
35%
30%
25%
20%
15%
10%
5%
0%
No Jumps
Large Jumps
Touch Pr
Default Pr

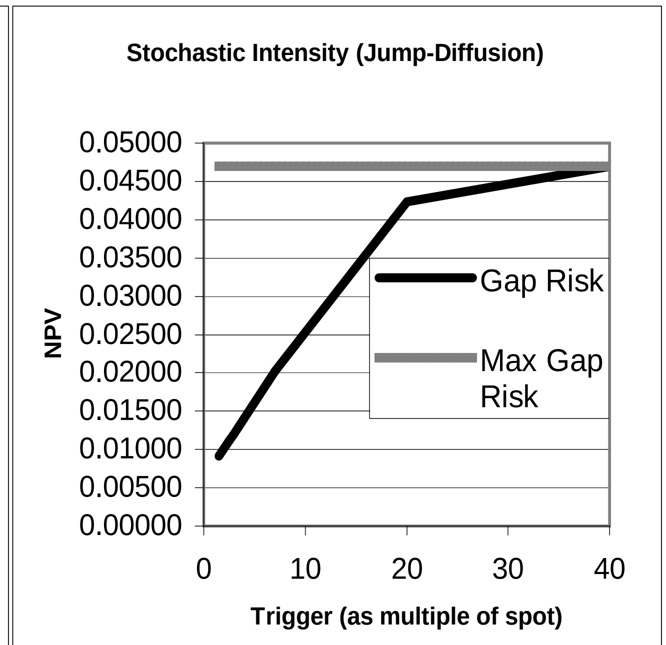

Salvation over all defaults
100%
90%
80%
70%
60%
50%
40%
30%
20%
10%
0%
Salv. Pr.
Salv. Pr.
No Jumps
Large Jumps


Salvation over all trigger touches
100%
80%
60%
40%
20%
0%
Salv. Pr.
Salv. Pr.
No Jumps
Large Jumps

Pricing a Note with a Trigger

The above trigger makes little sense according to the diffusive model, while it is an important protection according to the jump model. See that also the two models give two very different sensitivities of Gap Risk to triggers.





Pricing a Note with a Trigger

In a diffusive model, the range of movements of the spread is so limited that we reach very soon levels of the trigger such that it is almost impossible for it to be touched, sending Gap risk to maximum value. In case of possible high spread jumps, instead, the sensitivity to the trigger has a smooth behaviour, and allows to use it to control the level of the gap risk and minimize it. In this setting the alternative between two different model approaches (Structural vs Intensity) has been translated into a dependency from a parameter. The parameter can be moved smoothly allowing a consistent interpolation between the two modelling alternatives, for valuation and quantification of model risk.

Actions to mitigate model risk

Let us abstract now from the above example. We suppose that the parameter γ that we have used to parameterize our class of models is $0 \leq \gamma \leq 1$, and that the price $\Pi_t^T(\gamma)$ of some derivative (with maturity T) increases with γ. Suppose that the model validation process reaches the conclusion that $\Pi_t^T(\gamma = 0.5)$ is a reasonable price and validates the model with $\gamma = 0.5$. At the same time, the validation process recognizes that there is residual uncertainty about the value of γ.

Now the bank finds a counterparty which is eager to pay $\Pi_t^T(\gamma = 0.75)$ to buy the derivative. The difference

$$\mathbb{R} = \Pi_t^T(0.75) - \Pi_t^T(0.5)$$

appears a day-1 profit for the salespeople and the traders that closed the deal. The model risk managers should decide that $\mathbb{R}$ is not recognized to the trader at day-1, but is used to create a *model capital reserve*, that will be released to the trader only along the life of the derivative, for example from time t_1 to time t_2, $t_1 < t_2 < T$ the trader will receive only

$$\mathbb{R}\frac{(t_2 - t_1)}{T} \tag{8}$$

Model Reserves

What is the rationale of such a provision? Without it, salesmen and traders would be strongly motivated to sell large notionals of this derivative to counterparties eager to pay $\Pi_t^T(0.75)$, getting an important day-1 profit. However, due to the model uncertainty stated by in validation, the management should worry that it is possible that there is no day-1 profit, but just that the price $\Pi_t^T(0.5)$ coming from the bank's model is wrong. This may be the only reason why the bank is able to close many deals on this derivative!

With the reserve, the motivation to trade is reduced, because salesmen and traders see no immediate profit but just a reserve. Their P&L will not be increased by $\mathbb{R}$, and it will generate no end-of-year bonus. For seeing all of this profit, they have to remain in the bank until T. But if they do this, they will see also the potential losses of this derivative in its future life.

If the bank wants to reduce further the incentive to traders, a rule can be chosen to release the reserve according to a law slower than the linear releasing set in (8). An even stronger discouragement to trading this derivative at a model risk would be to compute always a reserve based on the most conservative model, thus considering $\Pi_t^T(1)$ when the product is sold, $\Pi_t^T(0)$ when the product is bought.

Model Lines or position limits

Another action to mitgate model risk is the introduction of *model position limits* or *model lines*, that are analogous to the credit lines used to manage credit risk. If the model for pricing some derivatives is considered subject to high model risk, the bank can validate a reasonable model but set a limit to the exposure built through these derivatives.

A bank has to decide first the total model line - the maximum exposure to model uncertainty allowed. This is a management decision.

Then a bank has to compute the specific ‘add-on’ of a given deal priced with the uncertain model, namely how much a single deal contributes to filling the total line.

Model Lines or position limits

How can banks estimate the potential loss due to model uncertainty? For a product bought at $\Pi(\bar{\gamma}, T)$, an estimate is

$$Notional * \left[\Pi_t^T(\bar{\gamma}) - \Pi_t^T(0)\right]. \tag{9}$$

Notice that, as it is reasonable, model lines should not regard deal closed at the conservative price $\Pi_t^T(0)$.

If the error we expect on γ is no more than $\triangle\gamma$, one can use, as a first order approximation to (9),

$$Notional * \left.\frac{\partial \Pi_t^T(\gamma)}{\partial \gamma}\right|_{\gamma=\bar{\gamma}} \triangle \gamma. \tag{10}$$

But we may be uncertain on the relevance of the point $\bar{\gamma}$ where the sensitivity is computed.

Model Lines or position limits

For credit lines the exposure is often a quantile of the future mark-to-market. For a future date $T_i < T$, and a confidence level of 90%, the exposure computed today at t is the level K_i such that

$$K_i = \min\left\{ K |\, \Pr\left(\Pi_{T_i}^{T}\left(\gamma\right) \leq K\right) = 90\%\right\},$$

and then one can take

$$\bar{K} = \max_i K_i$$

We can replace the exposures (9) and (10) with alternative quantities where $K_i\left(\gamma\right)$ replaces $\Pi_t^T\left(\gamma\right)$, or keep (9) and (10) since model risk is not associated to uncertainty about future values, but about present prices.

Notice that, as it happens with credit lines, a new deal priced with the uncertain model may actually reduce the exposure rather than increase it, for example if the deal is on the other side of the market compared to the majority of the existing deals.

Model Revisions

Specify dates a which the validation must be revised:

1. **Periodic Revisions**: scheduled regularly, for example every year. After one year from the first validation, a lot of new evidence may be available, probably there will be more visibility of traded quotes, and the knowledge of the model and its limits will be higher. The market may also have evolved.
2. **Triggered Revisions**. The Stress-test step may have revealed that under particular market conditions a model may become unreliable. If one of these market conditions becomes material, the Validation must be redone, and the model may require adjustments as a consequence of this. The revision dates in this case will be determined by quantitative triggers on market observable, or also on product features (time-to-maturity changes in time, for example).

Approximations: validate and then monitor the risk

Here we describe the main approximations which are used in the market practice for interest rate derivatives, particularly when the Libor and Swap Market Models are used. These approximations are usually first validated against more time-consuming methods such as montecarlo simulation. We will see first the steps for such tests. We then point out that the sneakiest model risk arises from most approximation being based of features of the market that are considered permanent. At times, however, they break down.

We take as examples two popular approximations: the formula for swaption pricing with the Libor Market Model and the formula for the computation of Convexity Adjustments. The validity of both approximations, extensively used for years, has been challenged by recent market shifts. The third approximation is more recent, and has not been challenged by market shifts, but we show which market moves could invalidate it. These analysis give indication on how to stress-test approximations and introduce quantitative indicators to monitor their validity.

The Swaption Volatility Approximation in the Libor Market model

Libor Rates and Interest Rate Swaps

We have already seen why the **Forward Libor Rate** $F(t; T_{i-1}, T_i)$ with expiry (fixing time) T_{i-1} and maturity (payment time) T_i can be written as

$$F(t; T_{i-1}, T_i) = \frac{1}{\tau_i}\left[\frac{P(t, T_{i-1})}{P(t, T_i)} - 1\right].$$

and we have also seen why $F(t; T_{i-1}, T_i)$ is a martingale under Q^i, the T_i-forward probability measure associated to $P(t, T_i)$. The Libor Market Model (BGM) is based on this, and adds, in its basic formulation, a lognormality assumption. Formally, one defines a set $\{T_0, \ldots, T_M\}$ of expiry-maturity dates which is the tenor structure with the corresponding year fractions $\{\tau_0, \ldots, \tau_M\}$. It models the simply compounded forward rates $F_k(t) = F(t; T_{k-1}, T_k)$ resetting at T_{k-1} (expiry) and with maturity T_k. Q^k is the T_k-forward probability measure associated to $P(t, T_k)$. We know $F_k(t)$ is a martingale under Q^k.

The Libor Market Model

The Libor Market Model assumes for each $F_k(t)$ under Q^k:

$$\boxed{dF_k(t) = \sigma_k(t) F_k(t)\, dZ_k^k(t), \quad t \leq T_{k-1},}$$

where Z^k is an M-dimensional Wiener process, the instantaneous correlation is matrix ρ, and $\sigma_k(t)$ is the instantaneous volatility function.

The Libor Market Model is automatically and exactly calibrated to the term structure when we set $F_k(0)$ equal to the corresponding market quoted forward rate. Also calibration to caps is easy, since the distributional assumptions are consistent with those underlying the market Caplet Black Formula.

However, each of the caplets composing a cap depends on a single forward rate. What if a product depends jointly on the dynamics of a plurality of forward rates? In this case one selects one single pricing measure Q^i, and all rates involved must be modelled under Q^i. This requires computing the dynamics of F_k under Q^i, $i \neq k$. We obtain:

LMM dynamics of $F_k(t)$ under Q^i, $k \neq i$

- $i < k, \quad t \leq T_i$:

$$dF_k(t) = \sigma_k(t) F_k(t) \sum_{j=i+1}^{k} \frac{\rho_{k,j} \tau_j \sigma_j(t) F_j(t)}{1 + \tau_j F_j(t)} dt + \sigma_k(t) F_k(t) dZ_k^i(t)$$

- $i = k, \quad t \leq T_{k-1}$: $\qquad dF_k(t) = \sigma_k(t) F_k(t) dZ_k^i(t)$
- $i > k, \quad t \leq T_{k-1}$:

$$dF_k(t) = -\sigma_k(t) F_k(t) \sum_{j=k+1}^{i} \frac{\rho_{k,j} \tau_j \sigma_j(t) F_j(t)}{1 + \tau_j F_j(t)} dt + \sigma_k(t) F_k(t) dZ_k^i(t)$$

The Libor Market Model: Joint Dynamics

Thus when a plurality of forward rates are modelled jointly, all rates but one have a dynamics that depends also on the level of other forward rates. These dynamics capture the no-arbitrage relationships that link together the different bits of the term structure, $F(t; T_{k-1}, T_k)$, $k = 1, \ldots, M$. These dynamics are derived using Girsanov Theorem and Change of Numeraire for dynamics.

What happens when we want to apply the model to swaptions?

The Libor Market Model: Joint Dynamics

The price of a swap fixing first in T_α and paying last in T_β is $\textsc{Swap}\Pi_t =$

$$\sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i(F(t,T_{i-1},T_i)-K) = P(t,T_\alpha)-P(t,T_\beta)-\sum_{i=\alpha+1}^{\beta}\left[P(t,T_i)\tau_i K\right].$$

The Swap rate (fair K) is

$$S_{\alpha,\beta}(t) = \frac{P(t,T_\alpha) - P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i}.$$

allowing to write

$$\textsc{Swap}\Pi_t = \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i(S_{\alpha,\beta}(t) - K)$$

The Swaption Black Formula

Recall the payoff of a Swaption:

$$
\begin{aligned}
\textsc{Swaption}_t &= D(t,T_\alpha)\left(\textsc{Swap}\Pi^{\alpha,\beta}_{T\alpha}\right)^+ \\
&= D(t,T_\alpha)\left(\sum_{i=\alpha+1}^{\beta} P(T_\alpha,T_i)\tau_i(S_{\alpha,\beta}(T_\alpha)-K)\right)^+ . \\
&= D(t,T_\alpha)\sum_{i=\alpha+1}^{\beta} P(T_\alpha,T_i)\tau_i(S_{\alpha,\beta}(T_\alpha)-K)^+
\end{aligned}
$$

In the market swaptions are quoted with the following Black formula:

$$
\textsc{Swaption}\Pi_t = \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i \textsc{Black}\left(S_{\alpha,\beta}(t), K, \int_t^{T\alpha} \sigma_{\alpha,\beta}(t)^2\, dt\right),
$$

Can we find also in this case a Change of Measure that allows to justify a Black formula?

Yes, but we need to change to a new measure, the $Q^{\alpha,\beta}$ Swap Measure, associate with the annuity or PV01 numeraire $\sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i$. Then

$$\begin{aligned}
\textsc{Swaption}\Pi_t &= \mathbb{E}_t^Q\left[\frac{B(t)}{B(T_\alpha)}\sum_{i=\alpha+1}^{\beta} P(T_\alpha,T_i)\tau_i(S_{\alpha,\beta}(T_\alpha)-K)^+\right] \\
&= \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i\mathbb{E}_t^{\alpha,\beta}\left[(S_{\alpha,\beta}(T_\alpha)-K)^+\right]
\end{aligned}$$

Under $Q^{\alpha,\beta}$, the swap rate $S_{\alpha,\beta}(t)$ is a martingale, since, by definition

$$S_{\alpha,\beta}(t)=\frac{P(t,T_\alpha)-P(t,T_\beta)}{\sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i}.$$

If one assumes lognormality of the swap rate,

$$dS_{\alpha,\beta}(t)=\sigma_{\alpha,\beta}(t)S_{\alpha,\beta}(t)\ dW_t^{\alpha,\beta}$$

then swaptions are priced with Black Formula.

If lognormality of the swap rate (Black for swaptions) is consistent with lognormality of libor rates (LMM) we can use the Black formula in LMM, once the volatility of the swap rate has been computed from the LMM parameters. We can exploit that the Swap Rate is a function of Libor Rates:

$$
\begin{aligned}
\textsc{Swap}\Pi_t &= \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i(F_i(t)-K)=0,\\
\boxed{S_{\alpha,\beta}(t)} &= \frac{\sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i F_i(t)}{\sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i} = \frac{\sum_{i=\alpha+1}^{\beta} \dfrac{P(t,T_i)}{P(t,T_\alpha)}\tau_i F_i(t)}{\sum_{i=\alpha+1}^{\beta} \dfrac{P(t,T_i)}{P(t,T_\alpha)}\tau_i}\\
&= \boxed{\sum_{i=\alpha+1}^{\beta} w_i(t)\ F_i(t),}\\
w_i(t) &= \frac{\tau_i \prod_{j=\alpha+1}^{i} \dfrac{1}{1+\tau_j F_j(t)}}{\sum_{k=\alpha+1}^{\beta} \tau_k \prod_{j=\alpha+1}^{k} \dfrac{1}{1+\tau_j F_j(t)}}.
\end{aligned}
$$

Testing LMM-SMM Consistency

To check if lognormality of the swap rate (SMM) is consistent with lognormality of libor rates (LMM), one proceeds as follows (see Brigo and Mercurio (2001):

1) Assume F_k's are lognormal under Q^k (hypothesis of the LMM).

2) Change the measure to see their dynamics under $Q^{\alpha,\beta}$ (Girsanov)

3) See which dynamics they imply for $S_{\alpha,\beta}(t)$ under $Q^{\alpha,\beta}$ (Ito)

4) Check if the distribution of $S_{\alpha,\beta}(t)$ under $Q^{\alpha,\beta}$ is lognormal

   (hypothesis of the SMM).

The answer is NO...however

Lognormal vs LMM swap rate

With a large number of LMM realizations of the Swap rate at T_α, one can compute the numerical density, and then compare it with a lognormal density. We see below that, in normal market conditions (2001-2006), the two distributions are hardly distinguishable.

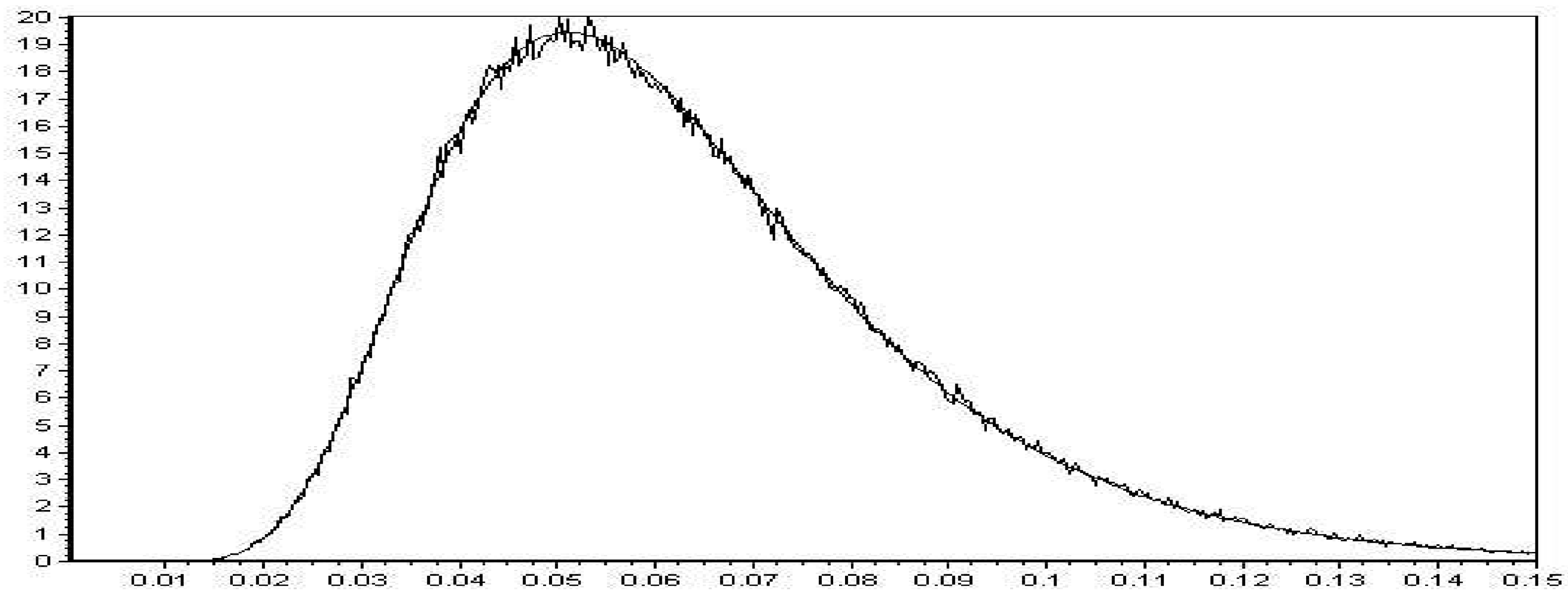


Figure 3: LMM Swap rate density from simulation vs a lognormal density.

Lognormal vs LMM swap rate

Treating the swap rate as lognormal even if we are in a model that assumes the Libor rates to be lognormal is an approximation that allows to use the Black formula.

$$\textsc{Swaption}\Pi_t = \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i \textsc{Black}\left(S_{\alpha,\beta}(t), K, \int_t^{T_\alpha} \sigma_{\alpha,\beta}(t)^2\, dt\right)$$

But we do not have the input for this formula, the volatility of the swap rate, which is a model parameter when one models directly the swap rate but in a LMM can only be computed as some function of the volatilities and correlations of Libor rates.

Approximations: what for?

When we have driftless lognormal dynamics

$$dX = \nu(t) X \, dW,$$

roughly one finds Black input V^2 as

1) $\dfrac{dX}{X} = \nu(t) \; dW$

2) $\dfrac{dX \, dX}{X^2} = \nu(t) \, \nu(t) \; dW \; dW = \nu(t)^2 \; dt$

3) $V^2 = \int_0^T \nu(t)^2 \; dt$

Approximations: what for?

Clearly, one can expect to find a deterministic number only when the dynamic is lognormal. Only in a lognormal model the *relative volatility* is deterministic. When dynamics is different, we may find at 2) a stochastic relative volatility

$$\tilde{\nu}\left(t, X\left(t, \omega\right)\right)^{2}\ dt.$$

Often, if there is a reason to assume the volatility of $\tilde{\nu}\left(t, X\left(t, \omega\right)\right)$ is small, the approximation for Black will be

$$\tilde{V}^{2} = \int_{0}^{T} \tilde{\nu}\left(t, \boxed{\ X\left(0\right)\ }\right)^{2}\ dt.$$

This is what we detail below for the swap rate.

Swaption Black Volatility

In the Swap Market Model, the swap rate is driftless lognormal under the swap measure

$$dS_{\alpha,\beta}(t) = \sigma_{\alpha,\beta}(t) S_{\alpha,\beta}(t) dW^{\alpha,\beta}.$$

with $\sigma_{\alpha,\beta}(t)$ deterministic instantaneous volatility. The volatility input of the Swaption Black formula is $\int_0^{T\alpha} \sigma_{\alpha,\beta}(t)^2 dt$, that one can compute as

$$\int_0^{T\alpha} \frac{dS_{\alpha,\beta}(t)\ dS_{\alpha,\beta}(t)}{S_{\alpha,\beta}(t)^2} = \int_0^{T\alpha} \frac{S_{\alpha,\beta}(t)^2 \sigma_{\alpha,\beta}(t)^2 dt}{S_{\alpha,\beta}(t)^2}$$

$$= \int_0^{T\alpha} \sigma_{\alpha,\beta}(t)^2 dt$$

How can we find an approximation for $\sigma_{\alpha,\beta}(t)$ in the Libor Market Model?

Industry LMM Approximation for Swap Vol, by Rebonato 98

The following approximations (Rebonato and Jaeckel 1998) come from analysis of the variability of the approximated objects. Start from swap rates expressed in terms of Libor

$$S_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} w_i(t) F_i(t) \cong \sum_{i=\alpha+1}^{\beta} w_i(0) F_i(t),$$

First Approx.

Thus the dynamics of the swap rate in the Libor Market Model is

$$dS_{\alpha,\beta}(t) = \sum_{i=\alpha+1}^{\beta} w_i(0)\ dF_i(t) = \sum_{i=\alpha+1}^{\beta} w_i(0)\ F_i(t)\sigma_i(t) dZ_i(t)$$

leading to

$$dS_{\alpha,\beta}(t)\, dS_{\alpha,\beta}(t) \ \cong \ \sum_{i=\alpha+1}^{\beta} \sum_{j=\alpha+1}^{\beta} w_i(0) w_j(0) F_i(t) F_j(t) \rho_{i,j} \sigma_i(t) \sigma_j(t)\, dt,$$

Industry LMM Approximation for Swap Vol

$$\sigma_{\alpha,\beta}(t) = \frac{dS_{\alpha,\beta}(t)dS_{\alpha,\beta}(t)}{S_{\alpha,\beta}(t)^2} \cong \sum_{i=\alpha+1}^{\beta}\sum_{j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(t)F_j(t)\rho_{i,j}\sigma_i(t)\sigma_j(t)}{\left(\sum_{i=\alpha+1}^{\beta} w_i(0)\ F_i(t)\right)^2} dt$$

$$\cong \sum_{i=\alpha+1}^{\beta}\sum_{j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}\sigma_i(t)\sigma_j(t)}{\left(\sum_{i=\alpha+1}^{\beta} w_i(0)\ F_i(0)\right)^2} dt$$

$\boxed{\text{Second Approx.}}$

so in the end the LMM Black volatility for swaptions is

$$\int_0^{T\alpha} \sigma_{\alpha,\beta}(t)^2 dt \cong$$

$$\boxed{\sum_{i=\alpha+1}^{\beta}\sum_{j=\alpha+1}^{\beta} \frac{w_i(0)w_j(0)F_i(0)F_j(0)\rho_{i,j}}{S_{\alpha,\beta}(0)^2} \int_0^{T\alpha} \sigma_i(t)\sigma_j(t)\, dt} = V_{\alpha,\beta}^{LMM}(T_\alpha)$$

Industry LMM Approximation for Swap Vol

Thus, in a Swap Market Model, swaptions are priced as

$$\textsc{Swaption}\Pi_t = \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i \textsc{Black}\left(S_{\alpha,\beta}(t), K, \int_t^{T_\alpha} \sigma_{\alpha,\beta}(t)^2 \, dt \right)$$

while in a Libor Market Model, where libor rates are the modelling variables, we have the following approximation

$$\begin{aligned} \textsc{Swaption}\Pi_t &= \sum_{i=\alpha+1}^{\beta} P(t,T_i)\tau_i \textsc{Black}\left(S_{\alpha,\beta}(t), K, V_{\alpha,\beta}^{LMM}(T_\alpha) \right), \\ V_{\alpha,\beta}^{LMM}(T_\alpha) &= \sum_{i=\alpha+1}^{\beta} \sum_{j=\alpha+1}^{\beta} \frac{w_i(0) w_j(0) F_i(0) F_j(0) \rho_{i,j}}{S_{\alpha,\beta}(0)^2} \int_0^{T_\alpha} \sigma_i(t)\sigma_j(t)\, dt \end{aligned}$$

.

Testing the accuracy of the LMM swap volatility formula vs MC

On 2002 data, we price the same swaptions with the same LMM model, but first we do it by montecarlo (that involves discretization but no approximation) and then by the above Black formula with input $V_{\alpha,\beta}^{LMM}(T_\alpha)$ (derived through different approximations). We will trust the Black formula approximation is reliable if $V_{\alpha,\beta}^{LMM}(T_\alpha)$ falls within the montecarlo window defined around the montecarlo implied volatility, once the montecarlo window has been reduced to a size comparable to the swaption bid-ask spread, and at the same time remains associated to a sufficiently high level of confidence.

MC Tests

In MC pricing of the payoff Π we estimate the price $\mathbb{E}(\Pi)$ as

$$\frac{\sum_{j=1}^{n} \Pi^j}{n},$$

where j goes across the n scenarios.

By Central Limit Theorem, for large n we have

$$\left(\frac{\sum_{j=1}^{n} \Pi^j}{n} - \mathbb{E}(\Pi)\right) \Big/ \left(\frac{Std(\Pi)}{\sqrt{n}}\right) \quad \sim \quad \mathcal{N}(0,1),$$

$$\begin{aligned} \Pr\left\{\left|\frac{\sum_{j=1}^{n} \Pi^j}{n} - \mathbb{E}(\Pi)\right| < \epsilon\right\} &= \Pr\left\{|\mathcal{N}(0,1)| < \epsilon\,\frac{\sqrt{n}}{Std(\Pi)}\right\} \\ &= 2\Phi\left(\epsilon\,\frac{\sqrt{n}}{Std(\Pi)}\right) - 1. \end{aligned}$$

MC Tests

We set first the $\Pr\{\}$.we want. Choosing $\Pr = 98\%$, and considering

$$2\Phi(z) - 1 = 0.98 \iff \Phi(z) = 0.99 \iff z \approx 2.33,$$

we have that the size ϵ of the window is

$$\epsilon \; = \; 2.33 \, \frac{Std(\Pi)}{\sqrt{n}},$$

thus we will raise n sufficiently to have ϵ not larger than the bid-ask. The final MC window is

$$\left[\frac{\sum_{j=1}^{n} \Pi^j}{n} - 2.33 \, \frac{Std(\Pi)}{\sqrt{n}}, \quad \frac{\sum_{j=1}^{n} \Pi^j}{n} + 2.33 \, \frac{Std(\Pi)}{\sqrt{n}} \right].$$

Testing the accuracy of the LMM swap volatility formula vs MC

Testing a 5×6 Swaption with calibrated volatilities and historical correlations:

MC volatility	MC inf	MC sup	Approximation
0.108612	0.108112	0.109112	0.109000

Stress test: 5×6 Swaption with rates upwardly shifted by 2%:

MC volatility	MC inf	MC sup	Approximation
0.097716	0.097301	0.098130	0.098048

Testing the accuracy of the LMM swap volatility formula vs MC

Stress test: 5×6 Swaption with augmented volatilities:

	MC volatility	MC inf	MC sup	Approx.
Cal. vol's $\times 1.2$	0.130261	0.129644	0.130878	0.130800
Cal. vol's $\times 2.5$	0.300693	0.298988	0.302399	0.303820

Testing a 10×10 Swaption

	MC volatility	MC inf	MC sup	Approx.
Cal. vol's	0.094937	0.094512	0.095363	0.095000
Cal. vol's $\times 1.2$	0.113853	0.113333	0.114372	0.114000
Cal. vol's $\times 2.5$	0.273425	0.272064	0.274788	0.277094

Lognormal vs LMM swap rate

This is confirmed by the shape of the numerical distribution for stressed volatility

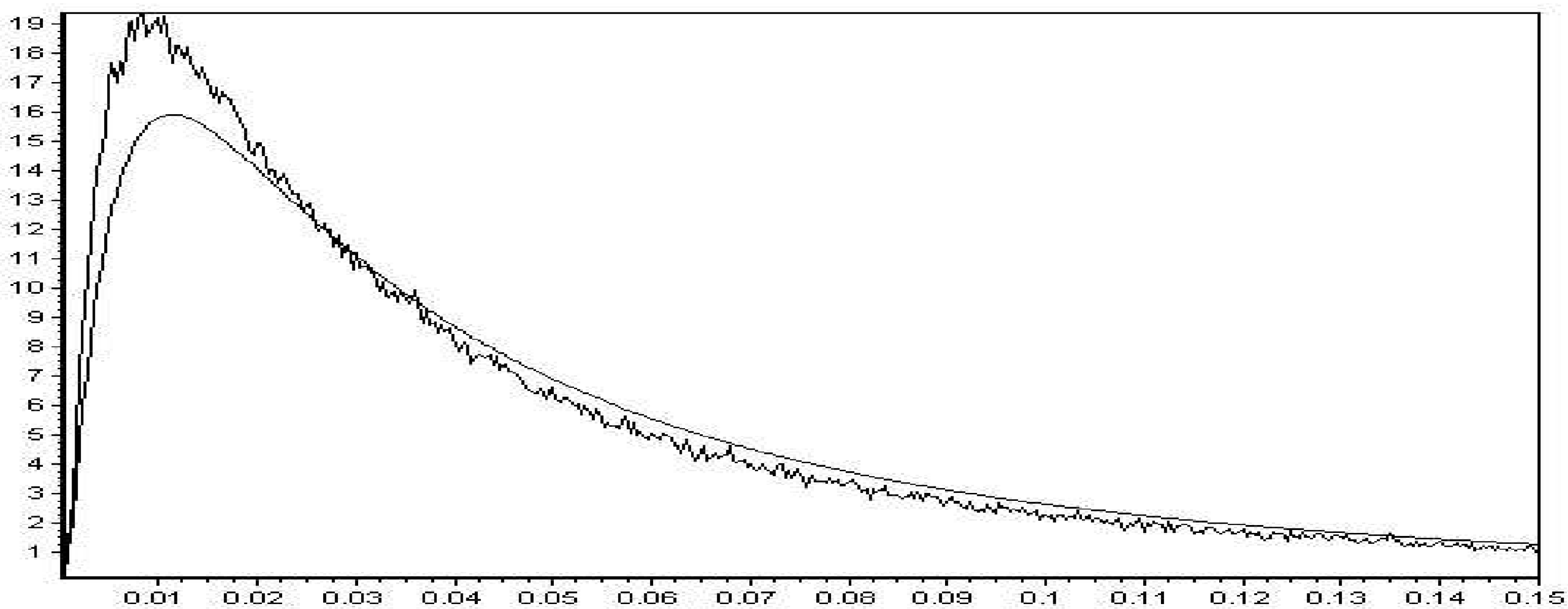


Figure 4: LMM Swap rate from simulation vs a lognormal density. Stress Case: market volatility doubled.

Swaption Volatility in the Crisis

The levels reached by swaptions in 2008-2009 rise a warning

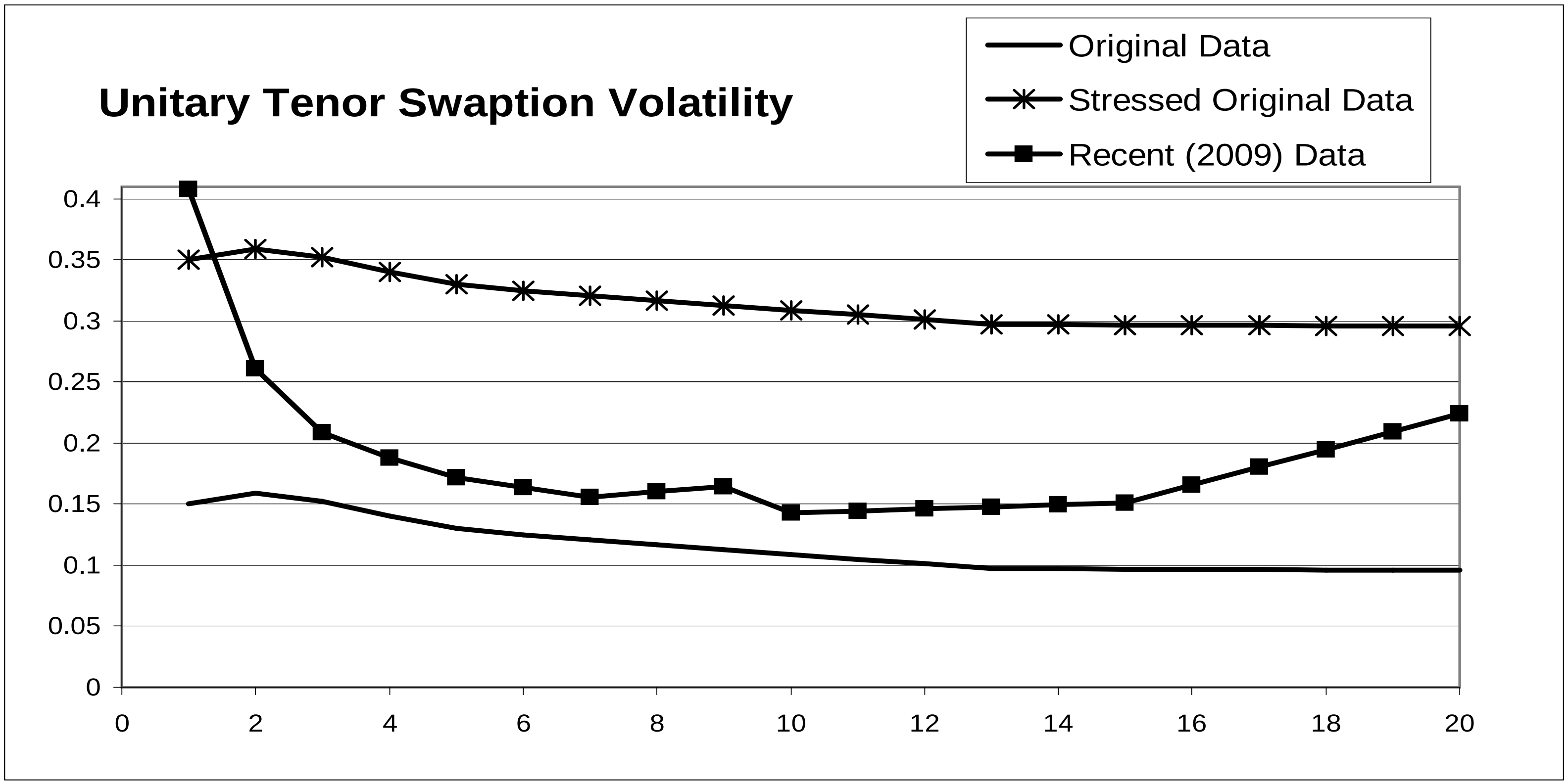


Figure 5: Unitary tenor swaption volatilities in January 2009 compared to volatilities used in previous stress tests.

The Market Approximation for CMS (Hagan)

Constant Maturity Swaps

Constant maturity swaps are contracts where one leg pays at T_i in

$$T_1, \ldots, T_n \text{ with } T_i - T_{i-1} = \delta$$

rate $S_{i\times c}(t)$, which is the swap rate with first reset in

$$\bar{T}_a = T_{i-1} \text{ (set in advance) or } \bar{T}_a = T_i \text{ (set in arrears)}$$

and paying at

$$\bar{T}_{a+1}, \bar{T}_{a+2}, \ldots, \bar{T}_{a+c} \text{ with } \bar{T}_j - \bar{T}_{j-1} = \tau$$

so that

$$S_{i\times c}(t) = S_{a,a+c}(t) = \frac{P(t,\bar{T}_a) - P(t,\bar{T}_{a+c})}{\sum_{j=a+1}^{a+c} P(t,\bar{T}_j)\tau}.$$

We consider first CMS in advance, so $S_{i\times c}$ sets in T_{i-1}, and define $\Delta := \frac{\delta}{\tau}$

Constant Maturity Swaps

The other leg pays Libor rates $F_i(t) = F(t, T_{i-1}, T_i)$ plus a spread $X_{n,c}$. The price of this contract (at $T_0 = 0$) is expressed as

$$
\begin{aligned}
& \mathbb{E}\left[\sum_{i=1}^{n} D(0,T_i)(S_{i\times c}(T_{i-1}) - F_i(T_{i-1}) - X_{n,c})\delta\right] \\
= \quad & \sum_{i=1}^{n} P(0,T_i)\left[S_{i\times c}(0) + \mathbf{CA}\right]\delta - \sum_{i=1}^{n}\left(P(0,T_{i-1}) - P(0,T_i)\right) - X_{n,c}\sum_{i=1}^{n} P(0,T_i)\delta \\
= \quad & \sum_{i=1}^{n} P(0,T_i)\left[S_{i\times c}(0) + \mathbf{CA}\right]\delta - \left(1 - P(0,T_n)\right) - X_{n,c}\sum_{i=1}^{n} P(0,T_i)\delta
\end{aligned}
$$

or equivalently

$$
= \sum_{i=1}^{n} P(0,T_i)\left(\left[S_{i\times c}(0) + \mathbf{CA}\right] - F_i(0) - X_{n,c}\right)\delta
$$

Constant Maturity Swaps

It is usually quoted via the so called CMS spread, the value of $X_{n,c}$ setting current price to zero:

$$X_{n,c} = \frac{\sum_{i=1}^{n} P(0,T_i)\left[S_{i\times c}(0) + \mathbf{CA}\right]}{\sum_{i=1}^{n} P(0,T_i)} - \frac{1 - P(0,T_n)}{\sum_{i=1}^{n} P(0,T_i)\delta}.$$

where $\mathbf{CA}$ is the **convexity adjustment**.

For the swap rate $S_{a\times c}(t)$ the classic convexity adjustment is

$$\mathbf{CA} \approx S_{a\times c}(0) \times \ \theta\left(S_{a\times c}(0)\right) \times \left(e^{\left(\sigma_{a\times c}^{\mathsf{ATM}}\right)^2 T_a} - 1\right), \tag{11}$$

$$\theta\left(S_{a\times c}(0)\right) = 1 - \frac{\tau S_{a\times c}(0)}{1 + \tau S_{a\times c}(0)}\left(\Delta + \frac{c}{\left(1 + \tau S_{a\times c}(0)\right)^c - 1}\right)$$

This approximation can be obtained via change of measure and the market model, as follows.

Convexity Adjustment

What does **CA** mean? Notice, applying basic change of numeraire, that we can write the price as

$$
\begin{aligned}
& \mathbb{E}\left[\sum_{i=1}^{n} D(0,T_i)(S_{i\times c}(T_{i-1}) - F_i(T_{i-1}) - X_{n,c})\delta\right] \\
= & \sum_{i=1}^{n} P(0,T_i)\left(\mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] - \mathbb{E}^{T_i}\left[F_i(T_{i-1})\right] - \mathbb{E}^{T_i}\left[X_{n,c}\right]\right)\delta \\
= & \sum_{i=1}^{n} P(0,T_i)\left(\mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] - F_i(0) - X_{n,c}\right)\delta.
\end{aligned}
$$

The Adjustment

Obviously the difficulty is to compute $\mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right]$. One can think of this quantity as

$$\begin{aligned}\mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] &= S_{i\times c}(0) + \mathbf{CA}, \\ \mathbf{CA} &= \mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] - S_{i\times c}(0) \\ \mathbf{CA} &= \mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] - \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\right]\end{aligned}$$

where $\mathbb{E}^{i\times c}$ is expectation under the swap measure associated with the numeraire $A^{i\times c}(t) = \sum P(t, T_i)\tau_i$, with summation over all life of the swap. How to compute $\mathbf{CA}$?

First of all, through change of numeraire, move also the Q^{T_i} expectation to the second measure under which $S_{i\times c}$ is a martingale.

$$\mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] = \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\frac{P(T_{i-1}, T_i)/A^{i\times c}(T_{i-1})}{P(0, T_i)/A^{i\times c}(0)}\right]$$

The Adjustment

$$
\begin{aligned}
\mathbf{CA} &= \mathbb{E}^{T_i}\left[S_{i\times c}(T_{i-1})\right] - \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\right] \\
&= \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\frac{P(T_{i-1},T_i)/A^{i\times c}(T_{i-1})}{P(0,T_i)/A^{i\times c}(0)}\right] - \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\right]
\end{aligned}
$$

Let us call

$$G_t = P(t,T_i)/A^{i\times c}(t),$$

so that

$$\mathbf{CA} = \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\left(\frac{G_{T_{i-1}}}{G_0} - 1\right)\right]$$

Some intuition on Convexity adjustment

Notice en passant that

$$G_t = P(t, T_i)/A^{i\times c}(t),$$

is a martingale under $Q^{i\times c}$ so

$$\begin{aligned}
\mathbf{CA} &= \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\left(\frac{G_{T_{i-1}}}{G_0} - 1\right)\right] \\
&= \frac{1}{G_0}\left\{\mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})G_{T_{i-1}}\right] - G_0 S_{i\times c}(0)\right\} \\
&= \frac{1}{G_0} Covariance\left(S_{i\times c}(T_{i-1}), G_{T_{i-1}}\right)
\end{aligned}$$

The Approximations

$$\mathbf{CA} = \mathbb{E}^{i\times c}\left[S_{i\times c}(T_{i-1})\left(\frac{G_{T_{i-1}}}{G_0} - 1\right)\right]$$

How to approximate this expression? We have to goals:
1) we would like also G_t to be expressed as a function of $S_{i\times c}(t)$.
2) we would like this function to be simple.

First we tackle 1). Expressing G as a function of S requires an approximation: that the forward swap rate S can be used a constant, flat rate for discounting all payments in the period spanned by the underlying swap. Thus discount is made by the factor $\frac{1}{(1+S\tau)}$ and, for example

$$P(t,T_i) = P(t,T_{i-1})\frac{1}{(1+S\tau)^{\Delta}},$$

First Approximation

while $A^{i\times c}(t)=\sum_{j=i}^{i-1+c}\tau P(t,\bar{T}_j)$ becomes

$$\begin{aligned}
A^{i\times c}(t) &= P(t,T_{i-1})\sum_{j=1}^{c}\frac{\tau}{(1+S\tau)^j}\\
&= P(t,T_{i-1})\frac{\tau}{(1+S\tau)}\left(\frac{1-\frac{1}{(1+S\tau)^c}}{1-\frac{1}{(1+S\tau)}}\right)\\
&= P(t,T_{i-1})\frac{1}{S}\left(1-\frac{1}{(1+S\tau)^c}\right).
\end{aligned}$$

First and Second Approximation

so

$$P(t,T_i)/A^{i\times c}(t) \approx G(S) = \frac{S}{(1+S\tau)^{\Delta}}\frac{1}{\left(1-\frac{1}{(1+S\tau)^{c}}\right)}$$

Two main simplifications:
A) the curve is assumed flat;
B) we are assuming perfect correlation between rates;

Now we consider our second goal, approximating $G\left(S_t\right)$ with a function as simple as possible. I can use the basic expansion

$$G\left(S_t\right) \simeq G\left(S_0\right) + G'\left(S_0\right)\left(S_t - S_0\right)$$

The Convexity Adjustment

Eventually, the adjustment is

$$\boxed{\mathbf{CA} \approx S_{i\times c}(0) \times \theta(S_{i\times c}(0)) \times \mathbb{E}^{i\times c}\left[\frac{S_{i\times c}(T_{i-1})^2}{S_{i\times c}(0)^2} - 1\right]}$$

With the alternative swap rate notation, $S_{a,b}(t)$, where a gives beginning of the swap and b gives the end,

$$\mathbf{CA} \approx S_{a,b}(0) \times \ \theta\left(S_{a,b}(0)\right) \times \mathbb{E}^{a,b}\left[\frac{S_{a,b}(T_a)^2}{S_{a,b}(0)^2} - 1\right], \tag{12}$$

where

$$\theta(S_{a,b}) := 1 - \frac{\tau S_{a,b}}{1+\tau S_{a,b}}\left(\Delta + \frac{b-a}{(1+\tau S_{a,b})^{b-a} - 1}\right)$$

The Convexity Adjustment

How to compute $E^{a,b}\left[S^2_{a,b}(T_a)\right]$? Assuming lognormality for the swap rate,

$$dS_{a,b}(t) = \sigma^{\mathsf{ATM}}_{a,b} S_{a,b}(t)\, dZ^{a,b}(t), \tag{13}$$

where $\sigma^{\mathsf{ATM}}_{a,b}$ is the at-the-money implied volatility for $S_{a,b}$, we have

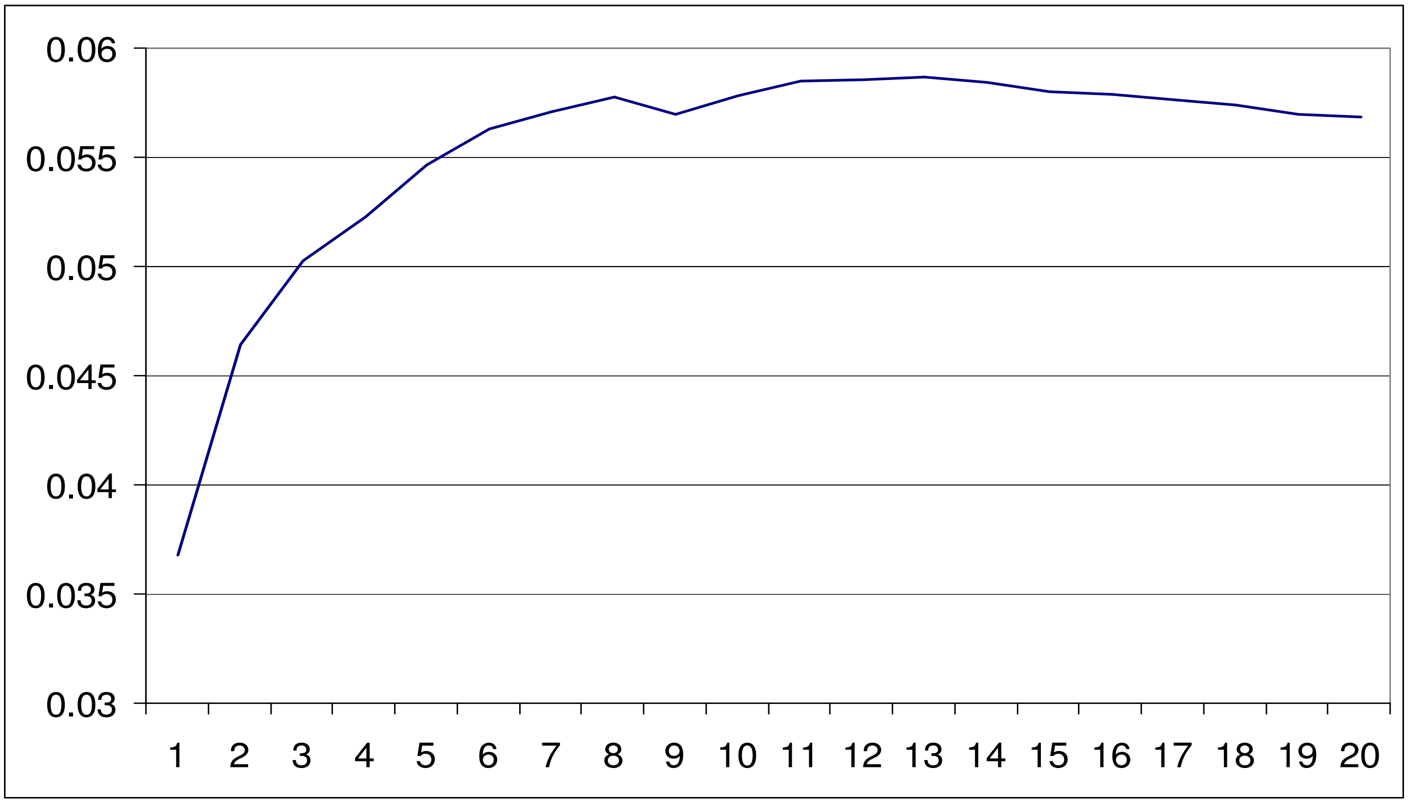
$$E^{a,b}\left[S^2_{a,b}(T_a)\right] = S^2_{a,b}(0)\, e^{(\sigma^{\mathsf{ATM}}_{a,b})^2\, T_a},$$

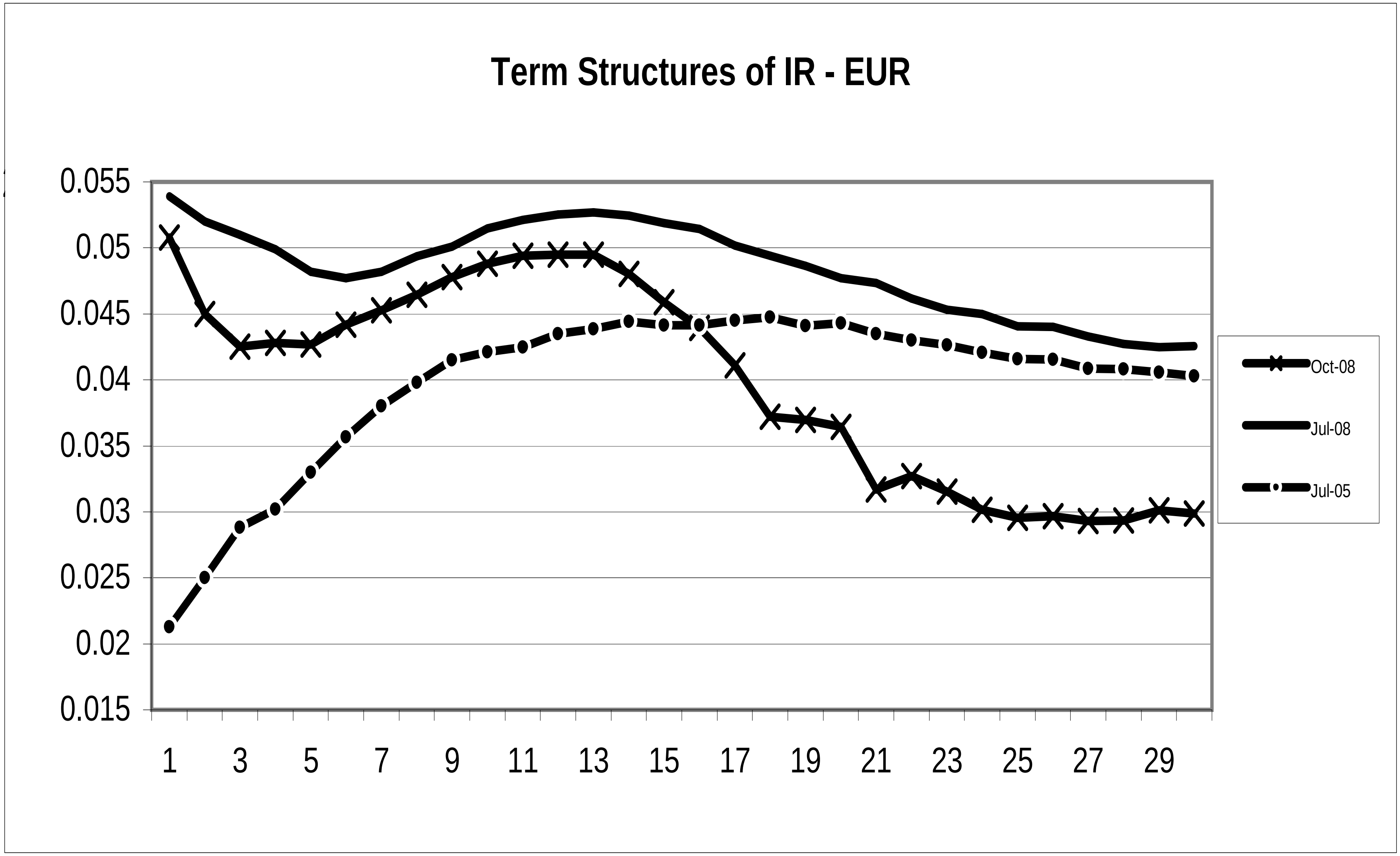
which leads to the classical convexity adjustment

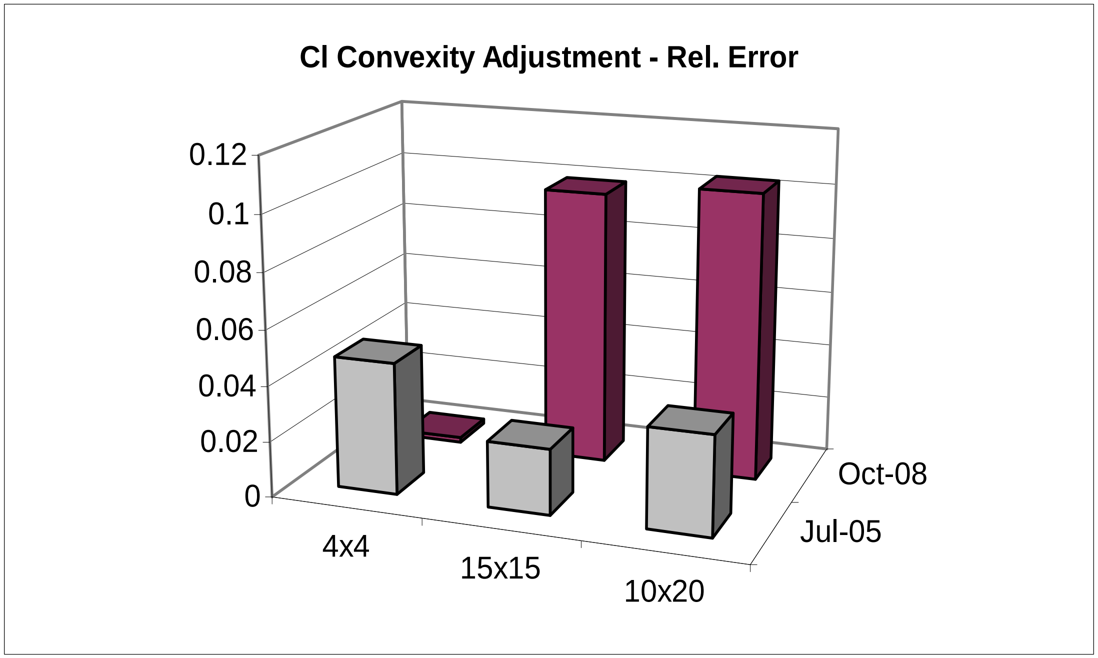
$$\boxed{\mathbf{CA} \approx S_{a,b}(0) \times \theta(S_{a,b}(0)) \times \left(e^{(\sigma^{\mathsf{ATM}}_{a,a+c})^2\, T_a} - 1\right)}.$$

PRICING CMS SWAPS: MONTECARLO, APPROXIMATIONS AND CONVEXITY ADJUSTMENTS

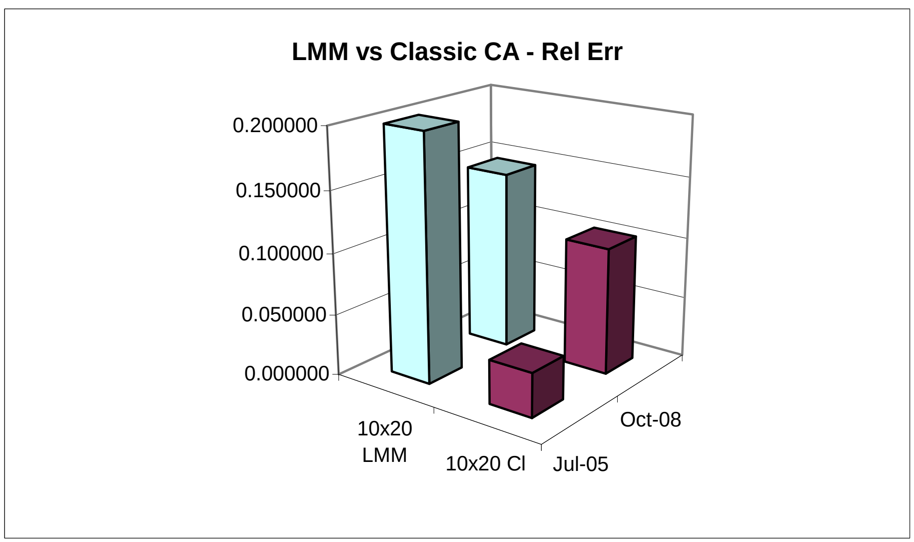
Maturity	2	2	2	6	6	6	10	10	10
Tenor	2	6	10	2	6	10	2	6	10
Swap Volatility	0.142000	0.122000	0.113000	0.116019	0.107059	0.101703	0.106000	0.098693	0.098000
MC Low	0.051377	0.054669	0.055913	0.057628	0.058373	0.058798	0.058889	0.059235	0.059406
MC Expect	**0.051343**	**0.054696**	**0.055937**	**0.057708**	**0.058441**	**0.058861**	**0.059000**	**0.059331**	**0.059502**
MC High	0.051308	0.054724	0.055961	0.057788	0.058509	0.058924	0.059110	0.059426	0.059599
MC Correction	**0.000174**	**0.000285**	**0.000376**	**0.000363**	**0.000783**	**0.001025**	**0.000571**	**0.001052**	**0.001605**
LMM Approx.	**0.051324**	**0.054708**	**0.055959**	**0.057745**	**0.058407**	**0.058903**	**0.059016**	**0.059402**	**0.059557**
%Err on Corr	**0.109195**	**-0.042105**	**-0.058511**	**-0.101928**	**0.043423**	**-0.040976**	**-0.028021**	**-0.067490**	**-0.034268**
Error	**0.000019**	**-0.000012**	**-0.000022**	**-0.000037**	**0.000034**	**-0.000042**	**-0.000016**	**-0.000071**	**-0.000055**
SMM Adjusted	**0.051321**	**0.054695**	**0.055943**	**0.057734**	**0.058405**	**0.058856**	**0.058999**	**0.059374**	**0.059506**
%Err Corr	**0.123705**	**0.004045**	**-0.017249**	**-0.070952**	**0.046412**	**0.004759**	**0.001931**	**-0.040912**	**-0.002620**
Error	**0.000022**	**0.000001**	**-0.000006**	**-0.000026**	**0.000036**	**0.000005**	**0.000001**	**-0.000043**	**-0.000004**


0.06
0.055
0.05
0.045
0.04
0.035
0.03
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20

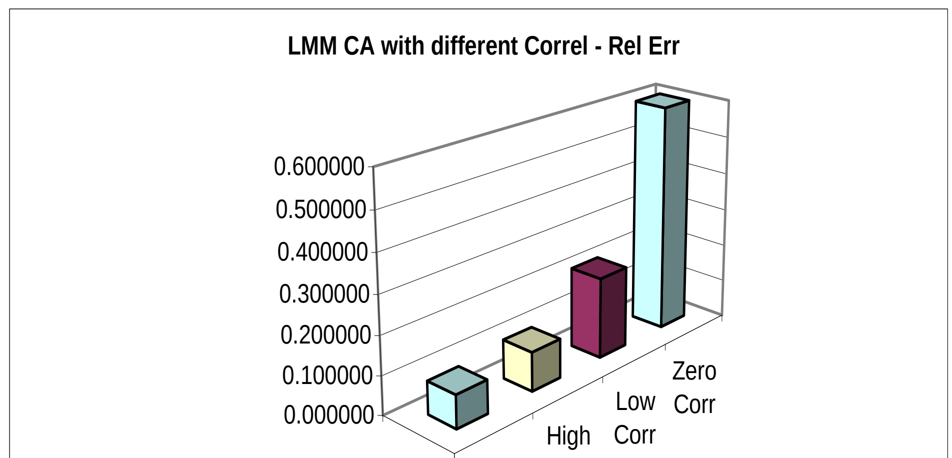

Term Structures of IR - EUR
0.055
0.05
0.045
0.04
0.035
0.03
0.025
0.02
0.015
1
3
5
7
9
11
13
15
17
19
21
23
25
27
29
Oct-08
Jul-08
Jul-05



Cl Convexity Adjustment - Rel. Error
0.12
0.1
0.08
0.06
0.04
0.02
0
4x4
15x15
10x20
Oct-08
Jul-05



LMM vs Classic CA - Rel Err
0.200000
0.150000
0.100000
0.050000
0.000000
10x20 LMM
10x20 Cl
Oct-08
Jul-05



LMM CA with different Correl - Rel Err
0.600000
0.500000
0.400000
0.300000
0.200000
0.100000
0.000000
Jul-05
Uni Corr
High Corr
Low Corr
Zero Corr

Exercise: why does it work?

In a standard LMM, the no-arbitrage drift of rates is usually approximated as follows:

$$\sum_{j=i+1}^{k} \frac{\tau_j \sigma_j \rho_{k,j} F_j(t)}{1+\tau_j F_j(t)} \approx \sum_{j=i+1}^{k} \frac{\tau_j \sigma_j \rho_{k,j} F_j(0)}{1+\tau_j F_j(0)}$$

This works for reason similar to the swaption approximation, and it has similar shortcomings. In a SABR LMM (Mercurio and Morini (2006)) the following approximation is used

$$\sum_{j=\gamma(t)+1}^{k} \frac{\tau_j \sigma_j \rho_{j,V} F_j^{\beta}(t)}{1+\tau_j F_j(t)} \approx \sum_{j=\gamma(t)+1}^{k} \frac{\tau_j \sigma_j \rho_{j,V} F_j^{\beta}(0)}{1+\tau_j F_j\,(0)}$$

1. Why is it more tricky than the standard one above?
2. Why in any case it works when the model is calibrated to market data?
3. When it could break?

The SABR Closed Form Formula

Here we see a different case, since we consider the comparison of an analytical formula, based on an approximation, with an equivalent one that includes no approximation but can be applied only in special case. The formula we consider is the SABR approximation for the implied volatility

$$dF(t) = V(t)F(t)^{\beta}\,dZ(t), \qquad dV(t) = vV(t)\,dW(t), \qquad V(0) = \alpha$$

$$\sigma^{\mathsf{SABR}}(K,F) = \frac{\alpha}{(FK)^{\frac{1-\beta}{2}}\left[1+\frac{(1-\beta)^2}{24}\ln^2\left(\frac{F}{K}\right)+\frac{(1-\beta)^4}{1920}\ln^4\left(\frac{F}{K}\right)+\cdots\right]}\frac{z}{x(z)}$$

$$\cdot\left\{1+\left[\frac{(1-\beta)^2\alpha^2}{24(FK)^{1-\beta}}+\frac{\rho\beta v\alpha}{4(FK)^{\frac{1-\beta}{2}}}+v^2\frac{2-3\rho^2}{24}\right]T_{k-1}+\cdots\right\},$$

$$z := \frac{v}{\alpha}(FK)^{\frac{1-\beta}{2}}\ln\left(\frac{F}{K}\right), \quad x(z) := \ln\left\{\frac{\sqrt{1-2\rho z+z^2}+z-\rho}{1-\rho}\right\}.$$

Validating the SABR Closed Form Formula

SABR embeds, for special parameters, the lognormal model and the constant elasticity of variance model, and therefore in these special cases the formula can be compared with alternative closed-form formulas. However in these cases one sets $\epsilon = 0$, which is not a very interesting case because the model is not anymore a stochastic volatility model, and because the formula comes from an expansion for small ϵ.

In the following we consider two peculiar parameterizations such that the price of a call option in SABR can be computed exactly in a very easy closed-form, but that allows considering any value for ϵ. It comes from an idea given to me by Bruno Dupire in Rome.

A little example on the inaccuracy of SABR formula

We assume $\rho = 1, \beta = 0$ so that

$$dF(t) = V(t)\, dW(t), \qquad dV(t) = \epsilon V(t)\, dW(t), \qquad V(0) = \alpha. \tag{14}$$

This implies

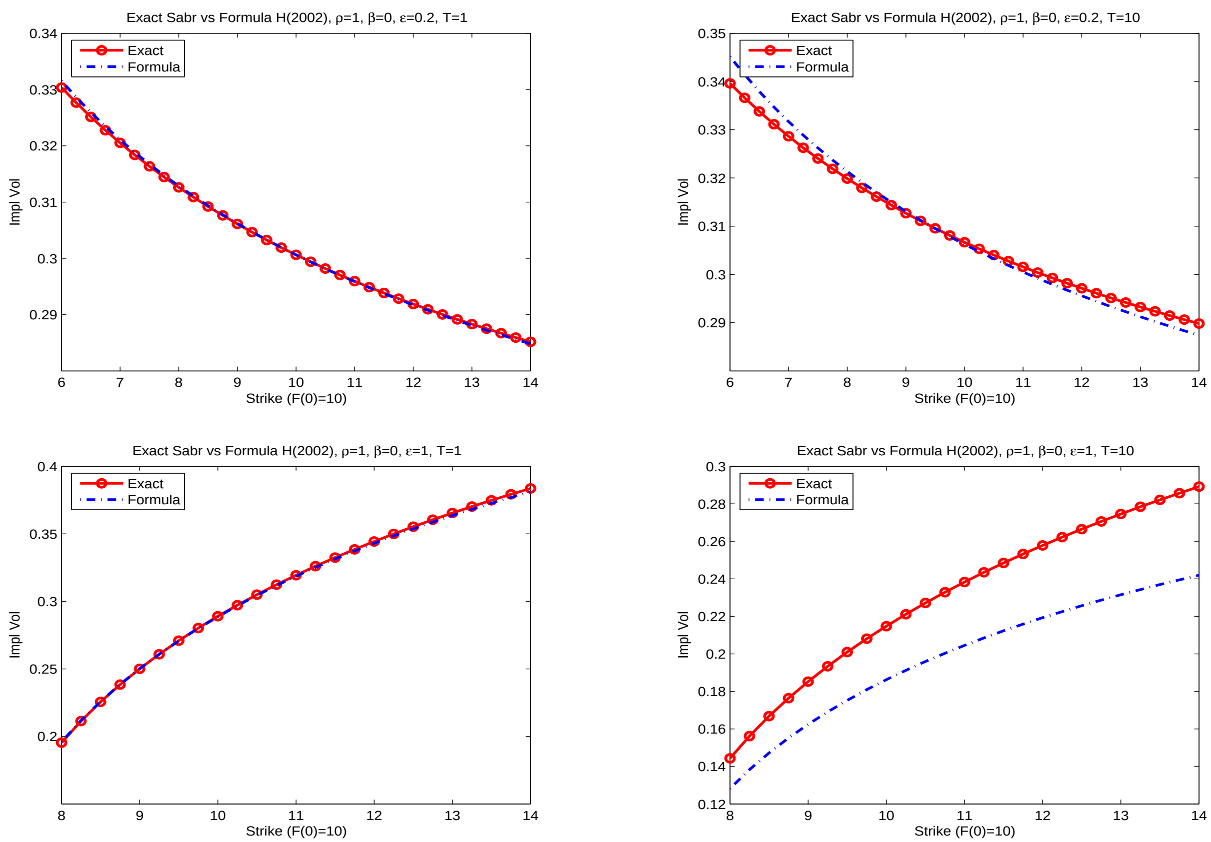
$$\begin{aligned} \frac{dV(t)}{\epsilon} &= V(t)dW(t) = dF(t), \\ F(T) &= F(0) + V(T)/\epsilon - \alpha/\epsilon. \end{aligned}$$

so that, by integration,where $V(t)$ is lognormal with instantaneous volatility ϵ. The price of a call with maturity T and strike K can be computed exactly:

$$\begin{aligned} \text{Call}(F, K, T) &= D(0,T)\, \mathbb{E}^Q \left[(F(T) - K)^+ \right] \\ &= \frac{1}{\epsilon} D(0,T)\, \mathbb{E}^Q \left[(V(T) - \alpha - K\epsilon + F(0)\epsilon)^+ \right] \\ &= \frac{1}{\epsilon} D(0,T)\, \text{Black}\left(\alpha, \alpha + K\epsilon - F(0)\epsilon, \epsilon\sqrt{T} \right), \end{aligned} \tag{15}$$

A little example on the inaccuracy of SABR formula

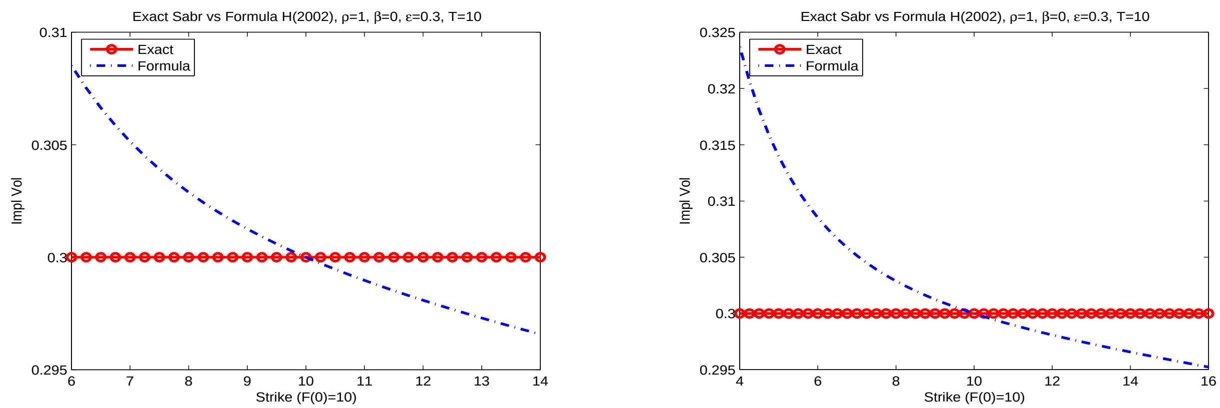
We consider call options when the underlying is $F(0) = 10$ and the (absolute) volatility is 30% of $F(0)$



The Hagan et al. (2002) formula appears very precise when the volatility of volatility is very low ($\epsilon = 0.1$) and the maturity is short (1 year). The precision worsens when we increase the maturity (10y).

A little example on the inaccuracy of SABR formula

Another interesting case is $\epsilon = 0.3$, since in this case one can prove that the model reduces to lognormal. In this case the results are as in the following figure.



Although in the latter case the level of the volatility of volatility is not high, the formula is not very precise, particularly if we extend the range of strikes (right pane in the above figure).

A little example on the inaccuracy of SABR formula

The second special case allowing for an exact closed-form formula considers $\rho = -1, \ \beta = 0$

$$dF(t) = V(t)\, dW(t), \quad dV(t) = -\epsilon V(t)\, dW(t), \quad V(0) = \alpha,$$

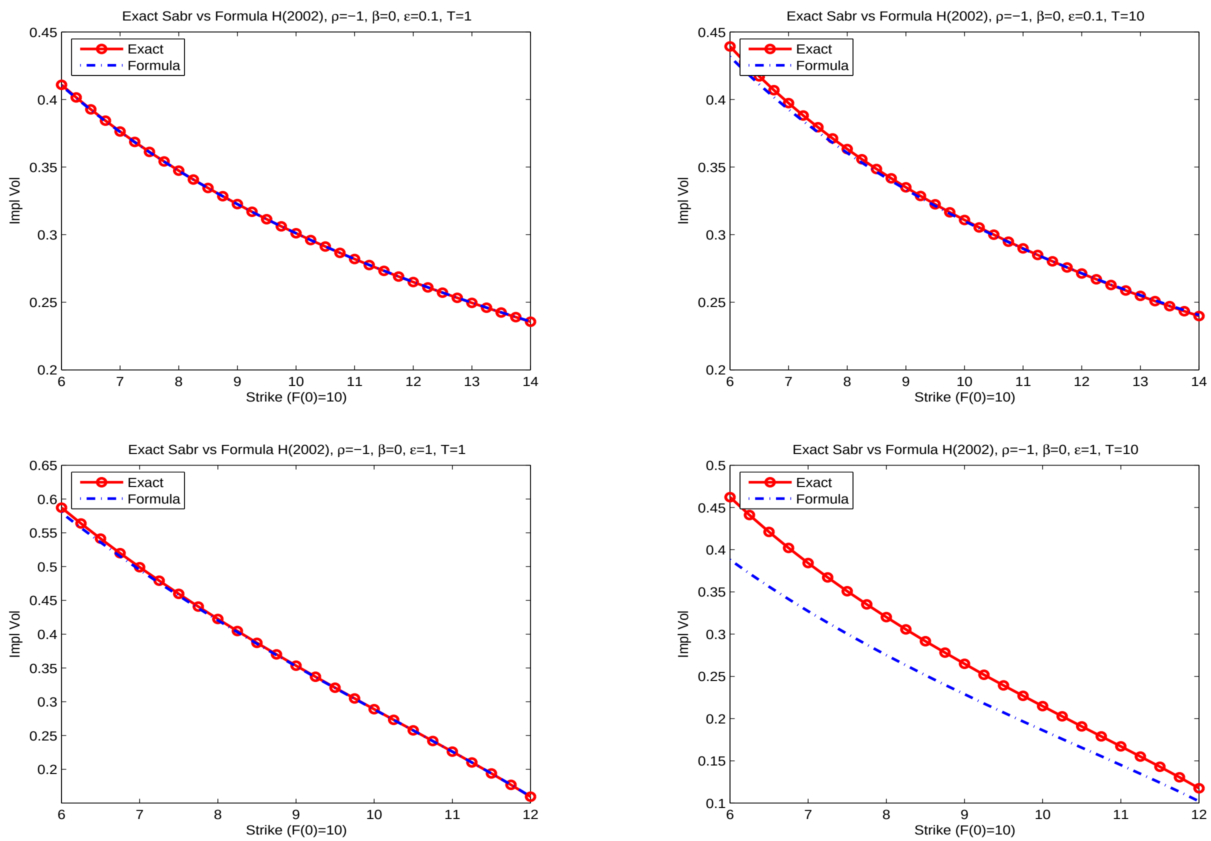
this implies

$$\begin{aligned} dV(t)/\epsilon &= -V(t)dW(t) = -dF(t), \\ F(T) &= F(0) - V(T)/\epsilon + \alpha/\epsilon \end{aligned}$$

where $V(t)$ is lognormal with instantaneous volatility ϵ. The price of a call with maturity T is

$$\begin{aligned} \textsc{Call}(F, K, T) &= \frac{1}{\epsilon} D(0.T)\, \mathbb{E}^Q \left[(-V(T) + \alpha - K\epsilon + F(0)\epsilon)^+ \right] \\ &= \frac{1}{\epsilon} D0, (T)\, \textsc{Black}^{Put} \left(\alpha, \alpha - K\epsilon + F(0)\epsilon, \epsilon\sqrt{T} \right), \end{aligned}$$

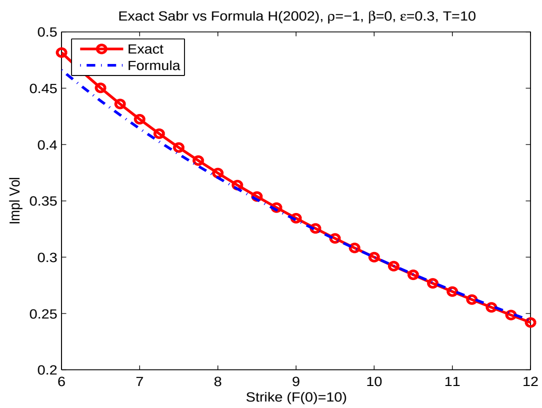
A little example on the inaccuracy of SABR formula

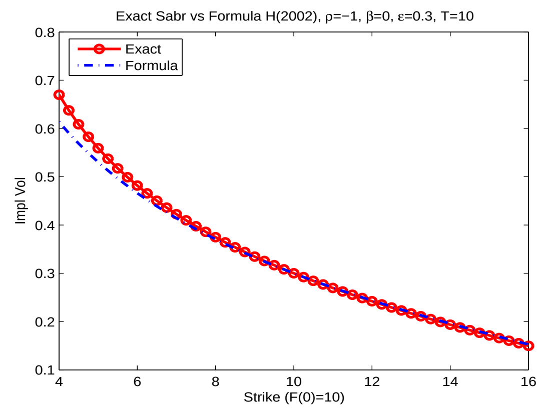


The results obtained for the $\rho = 1$ case appear confirmed also by the opposite $\rho = -1$ case.

A little example on the inaccuracy of SABR formula

We also show the case when $\epsilon = 0.3$





We finally mention that for $\rho = -1$ all configurations obtained correspond to decreasing skews similar to those that appear in the market, and the same applies to the $\rho = 1$ case for low or moderate volatility of volatility.

Stress-Testing Models and Stress Testing with models

The meaning of Stress-Testing

Stress testing is used in finance with two meanings.

PST *Portfolio stress-testing*. Stress-testing has been used a lot with reference to the assessments of the solidity of large banks performed by the FED in the US in 2009 and by the ECB and other European institutions in 2010 on European banks. Under this perspective stress-testing means assessing how a business (derivative, portfolio) would perform in an adverse financial scenario, in terms of predicted losses compared with capital and liquidity resources.

MST *Model-stress-testing*. Under this second meaning, stress-testing means subjecting the model's inputs to extreme values for understanding when a model breaks down, becoming imprecise or unstable. This is clearly relevant to model validation. This may regard a model based on an approximation that may become imprecise under some particular conditions. Various examples of this will be seen for interest rate models. In some other cases the solution remain mathematically correct under stress-conditions, but the model experiences a discontinuity and starts giving answers in contrast with logic and experience.

Portfolio Stress-Testing

How PST can be related to valuation models? The risk manager may design a scenario using personal judgement or some statistical model. The scenario, however, provides no value for the portfolio, one needs to translate the scenario into model inputs or parameters and then use the model to evaluate the portfolio.

The gaussian copula as used in the market is very difficult to use when the goal is to translate a financial scenario into model parameters. If stress testing becomes a very important component of risk management, *a model must also be suitable for stress testing*. Additionally a model that we cannot use to represent a stress scenario because we do not know how to do it is a model that is difficult to understand. Models with such features carry a high model risk.

To overcome the problems of the standard gaussian copula with flat correlation we introduce a different approach to correlations that improves dramatically in the joint fitting capability to the entire market skew with few meaningful parameters. This reveal elements on the market perception of risk that we would miss otherwise, so that meaningful stress-tests can be designed easily.

Portfolio Stress-Testing

We must be careful of the pitfalls. Expressing a financial scenario through a quantitative model is a delicate task. Models are difficult to understand when applied in a new domain. Our intuition can fail us and models can have a misleading behaviour, leading to consider worst-case scenarios that actually are best-case scenarios. We must be careful. Consider a tranche *not* spot starting priced with Gaussian Copula, we will find out that *only exceptionally the model is easy to stress*. Under general market conditions, the relation between its only parameter and the real-world risk of losses in the tranche is a very unstable relationship, so that detecting the real stress test is very difficult. We show three examples affected by this: the computation of the probability of concentrated losses like those the fear of which ignited the subprime crisis, the computation of dynamic credit VAR and the pricing of CDS counterparty risk.

This is a typical Model Stress-Test for variations of the payoff: we have found out that for forward-starting products an element of lack of realism partially negligible for spot-starting products becomes instead crucial. The stress test is in various directions, we first explore the product dimension, moving from spot-starting to forward-starting; then we explore the input dimension, both in the subdimension of the correlation parameters and in the one of the default intensity parameters.

Model Stress Testing. Designing Scenarios

Then another example for model stress-testing, now on market conditions rather than payoff feautures. We design different financial scenarios and we apply the model to them, to see if in all of these scenarios the model remains valid. we consider another aspect of the gaussian copula: the mapping methods used to give a correlation for a bespoke portfolio, particularly difficult to validate because for them we do not even have liquid quotes to assess their fitting capability or to perform a backtest on their validity. Thus we subject them to a stress test. How did we construct the financial scenarios? Using the data associated to important events in the past. For applying mapping to them we have to transform the historical data, by a projection from the time axis to the 'space axis', or the axis of different portfolios. This cannot guarantee that methods are right, but like a stress test it can tell us if they are wrong in some relevant cases. And it promptly detects a weakness which is then confirmed also by mathematical analysis. Having understood that standard methods see the relations between spreads and correlations only along one possible dimension (level of risk) while there has been historically another one (dispersion) that alternative methods can capture. we devise one such method that can be used as a benchmark for the quantification of model risk.

Examples on Copulas and Multiname Credit

We recall first a few technicalities about the CDO's and the copulas, then we see three examples that exemplify the aspects of stress-testing seen above. They also work as interesting examples of hidden model risk and model mistakes, and can provide some indication on how to design meaningful stress tests.

CDO and Tranched Loss

Protection Buyer	$\longrightarrow$ rate K at $T_{a+1}, \ldots, T_b$ on Outstanding Notional $\longrightarrow$	Protection Seller
	$\longleftarrow$ Tranched Loss Increment $dL_{A,B}(t)$ at defaults $\longleftarrow$	

The payoff of a CDO can be written as portfolio of options on the Loss of the pool, $L(t)$. In particular we write the loss of an equity tranche $[0, X]$ as

$$\begin{aligned} L_X(t) &= \frac{1}{X}\left[(L(t))\, 1_{\{L(t)\leq X\}} + X 1_{\{L(t)>X\}}\right] \\ &\quad \frac{1}{X}\left[L(t) - (L(t) - X)^+\right] \end{aligned}$$

and then each tranche $[A, B]$ can be written as the difference of two equity tranches, normalized by the size of the $[A, B]$ tranche

$$L_{A,B}(t) = \frac{1}{B-A}\left[B L_B(t) - A L_A(t)\right]$$

Linking defaults in intensity setting

In a standard intensity setting the default time τ_i of name i has the following distribution

$$\Pr\left(\tau_i \leq \tau\right) = F_{\tau_i}\left(\tau\right) = 1 - e^{-\int_0^{\mathcal{T}} \lambda_i(s) ds}. \tag{16}$$

We set $\int_0^{\tau} \lambda_i\left(s\right) ds = \Lambda_i\left(\tau\right) = \varepsilon_i$ and we know that, being $\Lambda\left(\cdot\right)$ increasing,

$$F_{\varepsilon_i}\left(\varepsilon\right) = F_{\tau}\left(\Lambda^{-1}\left(\varepsilon\right)\right) = 1 - e^{-\varepsilon},$$

so ε_i has a unit exponential distribution. The default times on n names can be written

$$\boldsymbol{\tau}_1 = \Lambda_1^{-1}\left(\varepsilon_1\right), \boldsymbol{\tau}_2 = \Lambda_2^{-1}\left(\varepsilon_2\right), \ldots, \boldsymbol{\tau}_n = \Lambda_n^{-1}\left(\varepsilon_n\right)$$

Or, under flat intensity,

$$\boldsymbol{\tau}_1 = \frac{\varepsilon_1}{\lambda_1}, \quad \boldsymbol{\tau}_2 = \frac{\varepsilon_2}{\lambda_2}, \quad \ldots, \quad \boldsymbol{\tau}_n = \frac{\varepsilon_n}{\lambda_n}$$

How can we relate defaults in this context?

Dependency in stochastic intensities

With stochastic processes for the intensities of two names, we can related the default times of the two names by correlating the intensities

$$
\begin{aligned}
d\lambda_i(t) &= \mu_i(t, \lambda_i(t))\, dt + \sigma_i(t, \lambda_i(t))\, dW_i(t) \\
d\lambda_j(t) &= \mu_j(t, \lambda_j(t))\, dt + \sigma_j(t, \lambda_j(t))\, dW_j(t) \\
dW_i(t)\, dW_j(t) &= \rho dt; \qquad \varepsilon_i \perp \varepsilon_j
\end{aligned}
$$

Advantages: possible tractability; ease of implementation; default of one name does not affect the intensity of other names (unrealistic but good for tractability). The correlation can be estimated historically from time series of credit spreads;

Disadvantages: unrealistically low dependence across $1_{\{\tau_j<t\}}$ and $1_{\{\tau_i<t\}}$, and between $\boldsymbol{\tau}_i$ and $\boldsymbol{\tau}_j$.

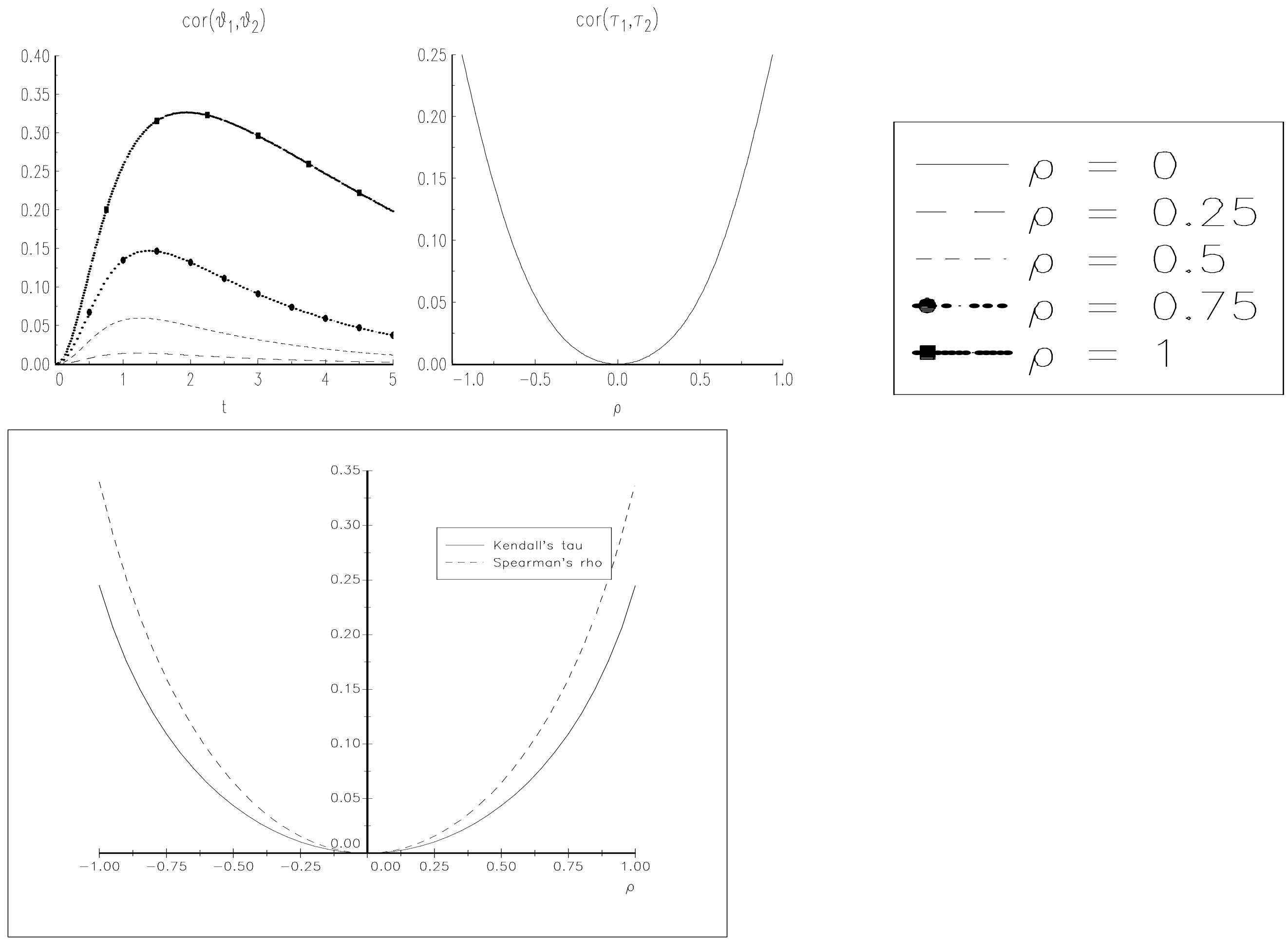


Figure 6: From Roncalli et al, using correlated quadratic intensities. Notice $\vartheta_i = 1_{\{\tau_i < t\}}$.

Threshold dependency

The market preferred to **set dependency among the ε_i of the different names and keep the** (stochastic or trivially deterministic) **intensities independent.** This is the framework that is currently used for correlation products in the market, especially for defining implied correlation.

Advantages: can take deterministic intensities, which makes life easier for the stripping of single name default probabilities; can reproduce sufficient levels of dependence across default times by putting dependence structures on the ε_i.

Disadvantages: we will see many when speaking of its main incarnation, the base correlation Gaussian copula... Let us now briefly recall it and see immediately some of its implications.

A general way of defining dependence: Copulas

The purpose of copulas is to separate information on the interdependence structure among the different variables from the information on the individual (marginal) distributions. Start from the joint distribution $F_{X_1,\ldots,X_n}$ of the variables $X = [X_1, \ldots, X_n]$:

$$F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = \Pr\left(X_1 \leq x_1, \ldots, X_n \leq x_n\right)$$

We set $u_i = F_{X_i}(x_i)$. Then the joint distribution $F_{X_1,\ldots,X_n}$ can be written as:

$$\begin{aligned} F_{X_1,\ldots,X_n}(x_1, \ldots, x_n) &= F_{X_1,\ldots,X_n}(F^{-1}_{X_1}\left(u_1\right), \ldots, F^{-1}_{X_n}\left(u_n\right)) \\ &= : C_{X_1,\ldots,X_n}(u_1, \ldots, u_n). \end{aligned} \tag{17}$$

The function $C_{X_1,\ldots,X_n}$ is the copula of the vector X.

Putting together Interdependence and Single names: COPULA

Notice: with this function $C_{X_1,\ldots,X_n}$, if we know the individual distributions F_{X_i} we can go back to the initial distribution.

$$C_{X_1,\ldots,X_n}(u_1,\ldots,u_n) = F_{X_1,\ldots,X_n}(F_{X_1}^{-1}(u_1),\ldots,F_{X_n}^{-1}(u_n))$$

Taking $u_i = F_{X_i}(x_i)$,

$$C_{X_1,\ldots,X_n}(F_{X_1}(x_1),\ldots,F_{X_n}(x_n)) = F_{X_1,\ldots,X_n}(x_1,\ldots,x_n)$$

So we **can reconstruct the Joint Distribution** from the **Copula** and the **Individual Distributions:**

$$\mathbf{F}_{1,2,\ldots,n}^{Joint} = \mathbf{Copula}(\mathbf{F}_1^{M\arg inal},\mathbf{F}_2^{M\arg inal},\ldots,\mathbf{F}_n^{M\arg inal}).$$

We have reached the goal of separating information on the dependency from information on the individual distributions. **The Copula contains all information on the dependency without information on the Individual distributions. Given the Copula, we can put the Individual distributions together to find the Joint distribution**.

Examples of copulas: Gaussian Copula

From the interdependence structure of the **standard** n-dimensional joint gaussian distribution Φ_{Σ}^{n}, with correlation matrix Σ,

$$\begin{aligned} & C_{\Sigma}^{\Phi}(u_1, \ldots, u_n) \\ = \quad & \Phi_{\Sigma}^{n}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)). \end{aligned}$$

This copula cannot be expressed in closed form. In the two dimensional case we have

$$C_{\rho}^{\Phi}(u, v) = \int_{-\infty}^{\Phi^{-1}(u)} \int_{-\infty}^{\Phi^{-1}(v)} \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{ \frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)} \right\} dsdt$$

No upper nor lower tail dependence for $\rho \neq 1$. The gaussian is easy to simulate, but its functional form is highly complex and overparameterized.

1-Factor Gaussian Copula

Define an n-variate Gaussian distribution as follows

$$X_i = \sqrt{\rho_i} M + \sqrt{1-\rho_i} Y_i \quad i = 1, \ldots, n \tag{18}$$

where Y_i, M are standard independent gaussian.

The distribution of X_i is Gaussian, with $\mathbb{E}[X_i] = 0$ and $Var[X_i] = \rho_i + 1 - \rho_i = 1$. Moreover

$$Corr(X_i X_j) = \sqrt{\rho_i \rho_j}$$

so the distribution is a simplified multivariate gaussian

$$\Pr(X_1 \le x_1, \ldots, X_n \le x_n) =: \Phi_{X_1,\ldots,X_n}(x_1, \ldots, x_n)$$

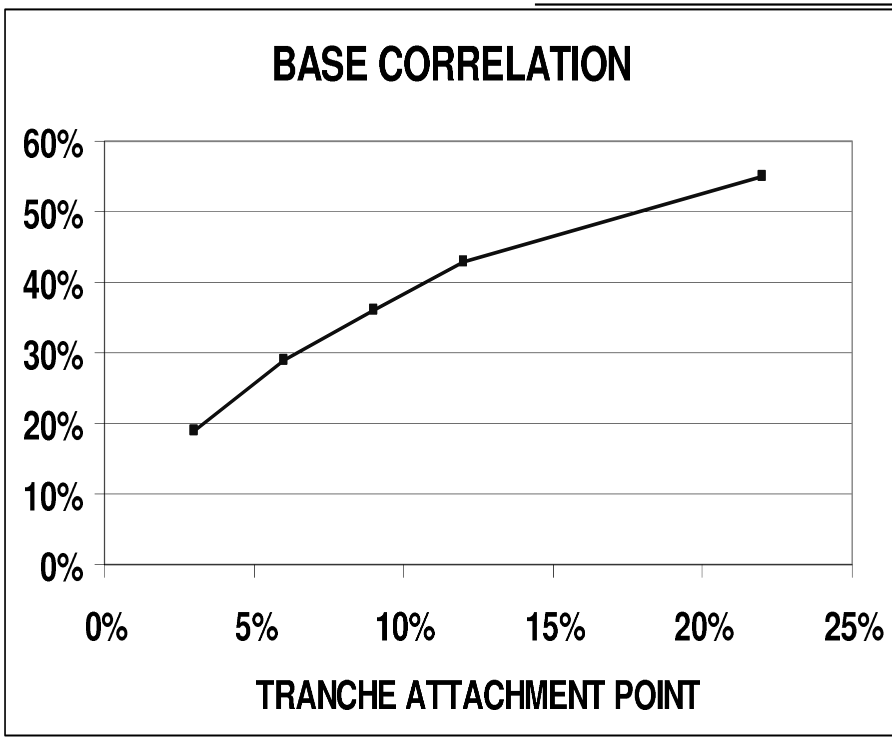
The copula is

$$C_{X_1,\ldots,X_n}(u_1, \ldots, u_n) = \Phi_{X_1,\ldots,X_n}(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n))$$

Flat correlation in gaussian copula: missing the link between correlation skew and systemic risk

Here we consider the relevance of having a model that can fit the market under the point of view of portfolio stress-testing. The capability of a model to fit liquid prices has nothing to do with the capability of the model to remain valid under stressed conditions. Getting the first-moment right does not imply capability to get the tails. However, here we point out a different aspect: models that do fit a large set of market observation with a consistent and parsimonious set of inputs give us insight about what is market perception of risk that we would miss otherwise, giving us a good indication about what a stress-test could be. We compare the standard tranche quotation model - Gaussian copula with Flat Correlation - with a different approach to correlation that allows to fit quotes in the tranche market.

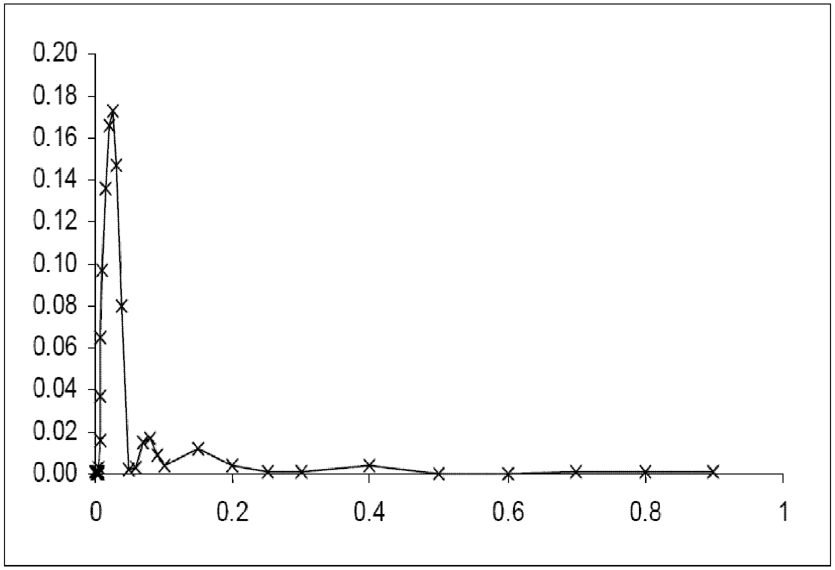
Correlation Skew



Gaussian Copula in the market is used with flat correlation: equal for all couples of names. This "model" can be fit to the CDO tranche market only by changing the correlation number for different tranches. This has been the market standard, but such one-tranche/one-parameter approach makes it difficult to design a consistent stress-scenario for the market

Scenario Density

There exists more complex models that fit the correlation skew with one single consistent set of parameters (first one is Hull and White's Perfect Copula) These models show that, after calibration to the correlation skew, the probability density of the Loss of the i-Traxx Index portfolio has some distinctive features, that we may call fat-tails or more precisely bumps in the tail.

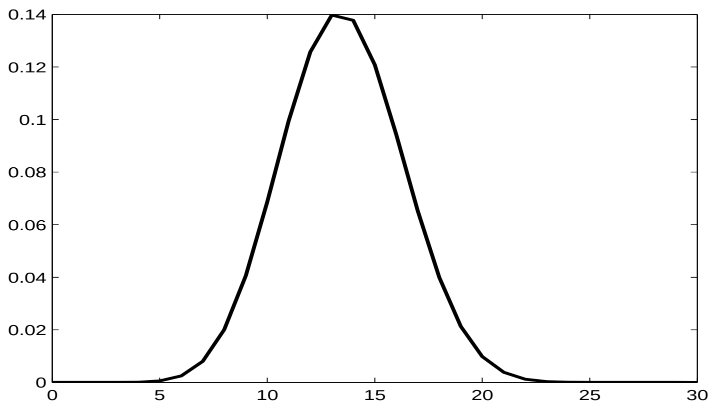


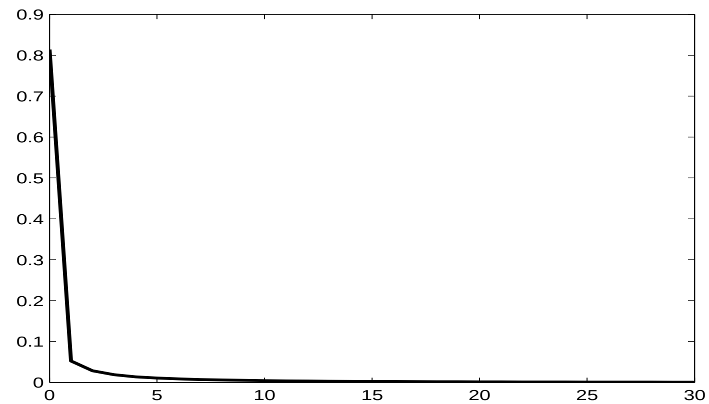
From Hull and White (2005) : Plot of the Index Loss density.

The Multimodal Loss Distribution: Fat Tails in Credit

The Loss density implied by market data is multimodal, with bumps in the far tail. This is a historical regularity. It is usually considered incompatible with gaussian copula.

Certainly, it is incompatible with the gaussian copula as it is implemented in market practice. In fact, with one single copula input, it is hard to model a realistic loss. See below two different cases, $\rho = 5\%$ $(\lambda = 0.15)$ and $\rho = 70\%$ $(\lambda = 0.01)$.





Can we incorporate multimodal losses, and therefore the correlation skew, sticking to the gaussian copula? We will see that in trying and doing this we will also learn one way of doing-stress test with this model for a portfolio of tranches.

An heterogeneous correlation structure

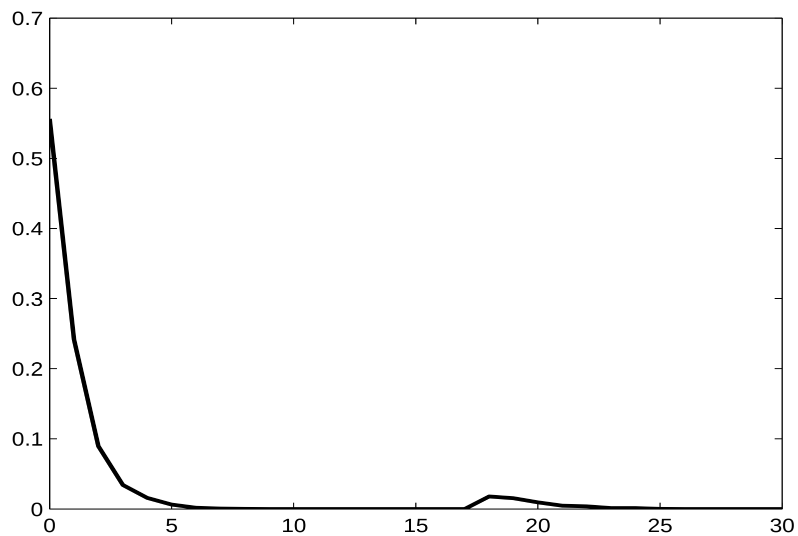
The main problems may lie in the way copula is used. Can we relate the existence of the correlation skew to the most strikingly unrealistic assumption in the market practice, namely flat correlation $\rho_{ij} = \rho$?

A model with a flat correlation around an average level tends to miss:
A) Some pairs of names are characterized by a **very high correlation** ($\rho_{ij} = 1$ in the most extreme case)
B) Some pairs of names are characterized by a **very low correlation** ($\rho_{ij} = 0$ in the most extreme case)

Notice that clusters of high correlations make **big losses and small losses** (far from average) **more likely**, while clusters of low correlations make **losses around the expected loss more likely**

Heterogeneous correlations

We now consider this issue in relation with the market multimodal loss distribution. We take a simple non homogeneous pool with two classes of names: a smaller group (40%) of riskier names, $\lambda = 0.015$, quite loosely related to any other (their risk is mostly idiosyncratic), and larger group of more senior names, $\lambda = 0.005$, that we expect to default only in case of a more generalized systemic (or sector-wide) crisis. Consistently we assign 20% correlation to the riskier names and unit correlation for the very senior names. The loss distribution is:



Heterogeneous correlations with systemic shock

The Gaussian copula with heterogeneous correlations, where a subset of senior names are assumed to be perfectly correlated, gives rise to clear multimodal loss densities, featuring bumps on the far tail. This is related to the fact that $\rho = 1$ is the only case when Gaussian copula features a tail dependency.

Can we go beyond this simple trick and use heterogenous correlation to fit the market? Let us see a simple solution.

Heterogeneous correlation as a function of credit risk

In the simple example, not only we used a heterogenous correlation, but we made correlation a function of default risk. This corresponds to the idea that names with different **levels** of credit risk are often associated to different **types** of credit risk: the risk of subinvestment grade borrowers usually comes from idiosyncratic, or firm-specific, risk factors, while the risk of senior borrowers usually comes from the risk of more systemic crisis.

The loans with the highest predictable losses have the lowest sensitivity to macroeconomic conditions. Conversely the lowest risk loans have the highest sensitivity to macroeconomic conditions (Breeden 2006).

This can be applied also to corporate senior names, including financials that, as we are seeing, are very sensitive to liquidity conditions in the global market. While there are a number of exceptions (fallen angels, business links), this appears a first ingredient to put in the correlation matrix: correlation must be a decreasing function of the names' default risk.

Heterogeneous correlation as a function of credit risk

We take the i-Traxx spreads as from February 2008, and we set us in the simplest possible gaussian correlation framework: One Factor Gaussian Copula. We define the n-variate Gaussian distribution as usual

$$\begin{aligned} X_i &= \rho_i M + \sqrt{1-\rho_i^2} Y_i \quad i = 1, \ldots, n \\ M, Y_i &\sim N(0,1), \qquad M \perp Y_i \perp Y_j, \quad i \neq j \end{aligned}$$

so that

$$\rho_{ij} = Corr(X_i X_j) = \rho_i \rho_j$$

However, we do not assume that the individual correlation parameters ρ_i are all the same, but instead we make them a function of the individual intensities or spreads:

$$\rho_i = f(Spread_i)$$

If ρ_i, ρ_j are high, then ρ_{ij} will also be high, while if either ρ_i, ρ_j is low the resulting correlation of the two names will be lower, and very low when both $\rho_i \approx 0$, $\rho_j \approx 0$.

Parameterizing Correlations

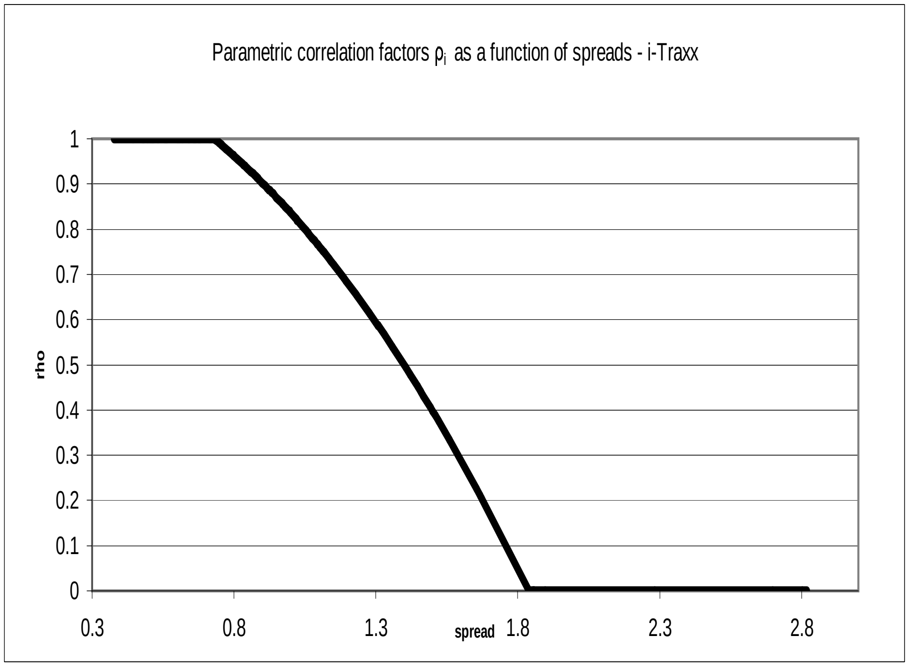
We can build a decreasing parameterization by setting

$$\rho_i \;=\; f\left(Spread_i\right) \quad > \quad \rho_j = f\left(Spread_j\right)$$
$$\text{when } \; Spread_i \quad < \quad Spread_j$$

This reminds of **local correlation**: models that make a flat correlation be an increasing function of the loss $L\left(t\right)$, so as to change it when approaching different detachments. Here we do it implicitly: when more risky names default through idiosyncratic events, and the loss approaches more senior tranches, the remaining names have a higher average correlation. However here we remain in a standard gaussian copula.

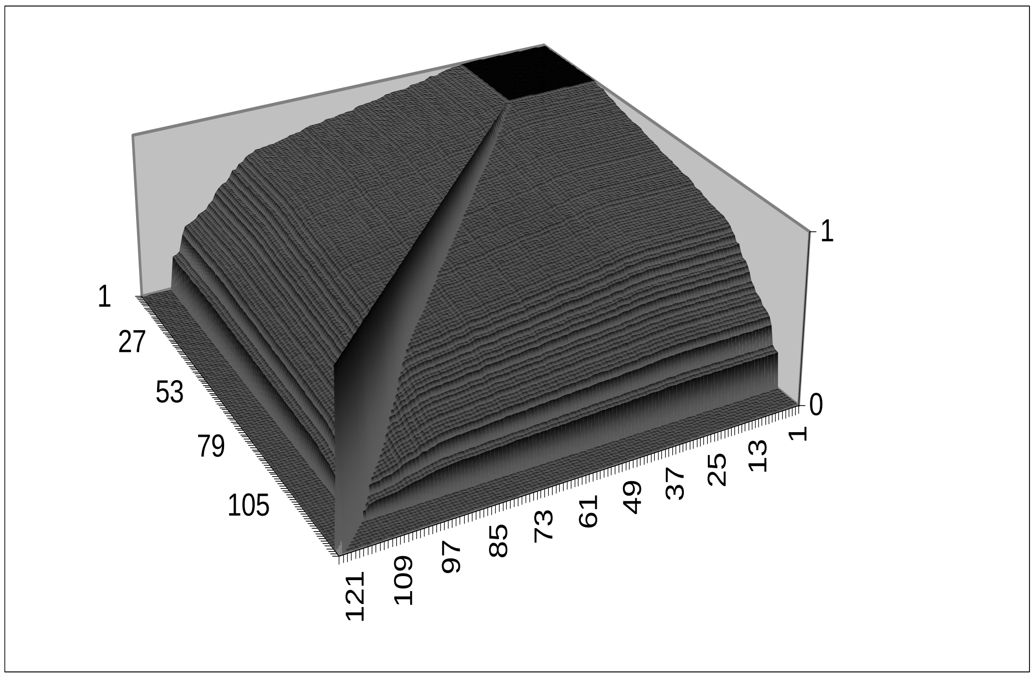
Parameterizing Correlations

We also want the parameterization to be sufficiently general to allow for correlation factors as high as $\rho_i = 1$ and as low as $\rho_i = 0$. A possible parameterization of this kind is a 3-parameter constrained polynomial (see Duminuco and Morini (2008)). On February 15, 2008, this results in the following dependence of the correlation factors ρ_i on the spreads of the i-Traxx names:



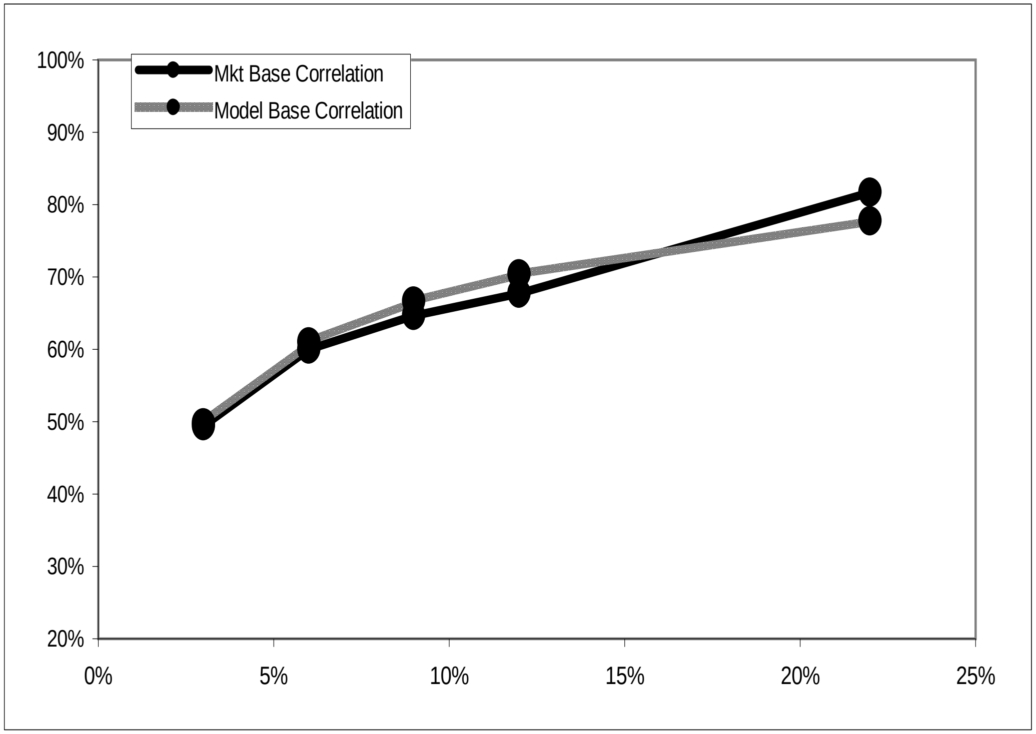
Parametric Correlation: the resulting Matrix

This parameterization corresponds to the heterogenous correlation:



Parametric Correlation: Implied Correlation Skew

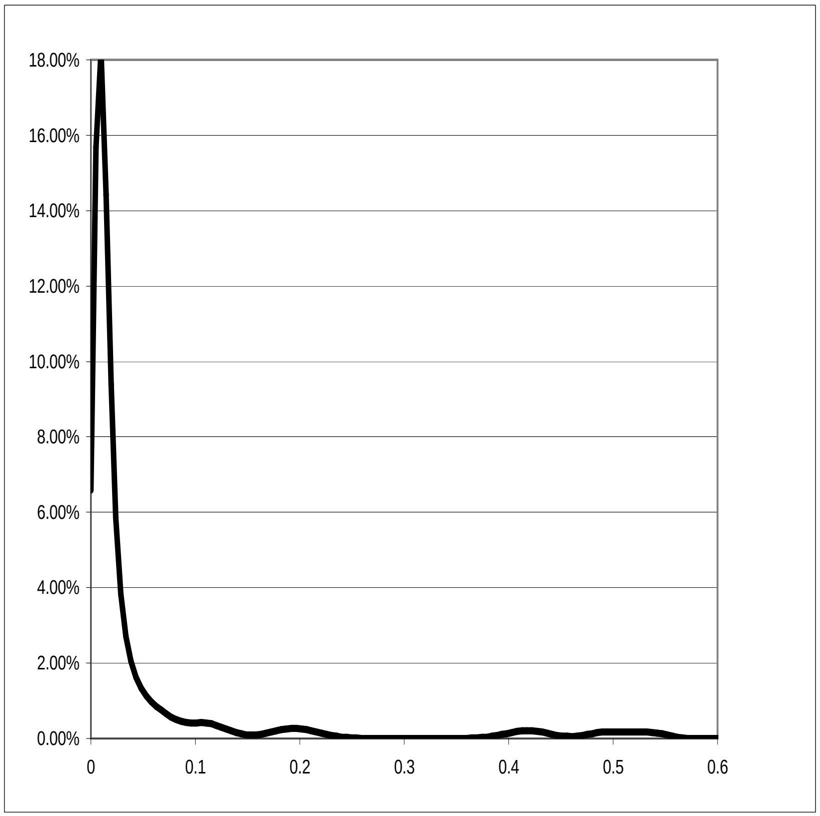
This correlation implies a rather realistic correlation skew, consistent with market correlation skew of February 15



One can take more factors (e.g. the financial sector factor) or scenarios, to increases the flexibility in fitting correlation skews.

Parametric Correlation: the implied Loss Distribution

The implied Loss distribution can be easily computed, and it has the desired multimodal structure, although we are still in a standard Gaussian Copula framework.



Parametric Correlation

With this approach to Gaussian Copula, not only can we fit the market, but also we have a framework that allows stress testing. The capability of this very simplified model to fit the market suggests that the traditional approach of rating agencies to set correlations historically based on only two dimensions - geographical proximity and industry sector - was letting out an important aspect of correlation risk.

A paradox in default correlation: missing concentration risk

The Gaussian Copula with deterministic default intensity (no spread volatility) has been deeply criticized after the credit crunch for a number of reasons: the correlation inputs were wrong (too low), the Gaussian Copula does not feature tail-dependency leading to a portfolio distribution with no fat-tails. Here we focus on a different problem: in some apparently standard applications the model has a misleading behaviour, that in particular **can lead us to make wrong stress-testing, using a worst-case scenario that actually is a best-case scenario**. The problem arises when one tries to apply the gaussian copula to the evaluation of the probability of the events that led to the credit crunch: concentration of losses *in time*.

Linking Defaults

In a standard intensity setting the default time τ_i of name i has the following distribution

$$\Pr\left(\tau_i \leq \tau\right) = F_{\tau_i}\left(\tau\right) = 1 - e^{-\int_0^{\mathcal{T}} \lambda_i(s)ds}. \tag{19}$$

We set $\int_0^{\tau} \lambda_i\left(s\right) ds = \Lambda_i\left(\tau\right) = \varepsilon_i$ and we know that, being $\Lambda\left(\cdot\right)$ increasing,

$$F_{\varepsilon_i}\left(\varepsilon\right) = F_{\tau}\left(\Lambda^{-1}\left(\varepsilon\right)\right) = 1 - e^{-\varepsilon},$$

so ε_i has a unit exponential distribution. The default times on n names can be written

$$\boldsymbol{\tau}_1 = \Lambda_1^{-1}\left(\varepsilon_1\right), \boldsymbol{\tau}_2 = \Lambda_2^{-1}\left(\varepsilon_2\right), \ldots, \boldsymbol{\tau}_n = \Lambda_n^{-1}\left(\varepsilon_n\right)$$

Or, under flat intensity,

$$\boldsymbol{\tau}_1 = \frac{\varepsilon_1}{\lambda_1}, \quad \boldsymbol{\tau}_2 = \frac{\varepsilon_2}{\lambda_2}, \quad \ldots, \quad \boldsymbol{\tau}_n = \frac{\varepsilon_n}{\lambda_n}$$

How can we relate defaults in this context?

Linking Defaults

$$\boldsymbol{\tau}_1 = \Lambda_1^{-1}(\varepsilon_1), \boldsymbol{\tau}_2 = \Lambda_2^{-1}(\varepsilon_2), \ldots, \boldsymbol{\tau}_n = \Lambda_n^{-1}(\varepsilon_n)$$

The market approach is to set dependency among the ε_i of the different names and keep the intensities deterministic. The dependency is introduced using a copula function, so that

$$\mathbf{F}_{1,2,\ldots,n}^{Joint} = \mathbf{Copula}(\mathbf{F}_1^{M\arg inal}, \mathbf{F}_2^{M\arg inal}, \ldots, \mathbf{F}_n^{M\arg inal}).$$

The copula separates information on the dependency from information on the individual distributions. The Copula contains all information on the dependency without information on the Individual distributions. Given the Copula, we can put the Individual distributions together to find the Joint distribution. There need to be no relation between individual distributions and Copula function. The market chose for default dependency the Gaussian Copula, that depends on a matrix of correlations ρ_{ij}.

Last to Default

In a portfolio of n names we indicate: $\tau^{(j)}$ =time of the j-th default in the portfolio. The payoff of a Last-to-Default contract is

Protection Buyer	$\rightarrow$	rate S at $T_{a+1}, \ldots, T_b$ before $\tau^{(n)}$	$\rightarrow$	Protection Seller
	$\leftarrow$	Protection $\mathrm{LGD}^{(n)}$ at $\tau^{(n)}$ if $T_a < \tau^{(n)} \leq T_b$	$\leftarrow$	

In a gaussian copula with deterministic intensity the value of the protection increases with correlation, since logically there is more probability of all names defaulting together in the same period of time when default correlation is higher. We also have an interesting result for the limit case where the default correlation among all names is set to $\rho = 1$:

- Correlation 1: S is the same as the spread paid on CDS written on the least risky name

On the other hand:

- Correlation 0: S is around the product of the individual spreads, usually very low

Last to Default

$$\boldsymbol{\tau}_1 = \Lambda_1^{-1}(\varepsilon_1), \boldsymbol{\tau}_2 = \Lambda_2^{-1}(\varepsilon_2), \ldots, \boldsymbol{\tau}_n = \Lambda_n^{-1}(\varepsilon_n)$$

Or, under flat intensity,

$$\boldsymbol{\tau}_1 = \frac{\varepsilon_1}{\lambda_1}, \quad \boldsymbol{\tau}_2 = \frac{\varepsilon_2}{\lambda_2}, \quad \ldots, \quad \boldsymbol{\tau}_n = \frac{\varepsilon_n}{\lambda_n}$$

if $\rho = 1$ all ε are the same ($\varepsilon = \varepsilon_1 = \varepsilon_2 = \ldots = \varepsilon_n$). while the individual models of default, expressed by the λ_i's, remain what they were before being joined together and can be very different from each other. When $\rho = 1$: $\boldsymbol{\tau}_i = \dfrac{\varepsilon}{\lambda_i}$, with the last default time (highest) being the one associated to lowest λ_i. Thus:

1) names with extremely different default risk can have unit default correlation, $\rho = 1$

2) when $\rho = 1$, such names can have defaults which are largely far apart in time.

Default correlation in the gaussian copula does not have the same temporal meaning that we usually give to the concept of default correlation. Can this be misleading in some relevant market cases, and create some risk of model misinterpretation? Let us see an example...

A paradox in market copula: Forward-start Last to Default

Consider we are exposed to losses from default of two names (or two different investments). We are not so much worried about the probability that they default, but we are very worried about the probability that they **both** default in a particular, **quite short period of time**, because we know that in that period we may run short of liquidity...It does not sound such a bizarre worry, after the credit crunch.

For example, we may be worried about the period between 3y and 5y from now, and we are exposed to two investments with the following default intensities:

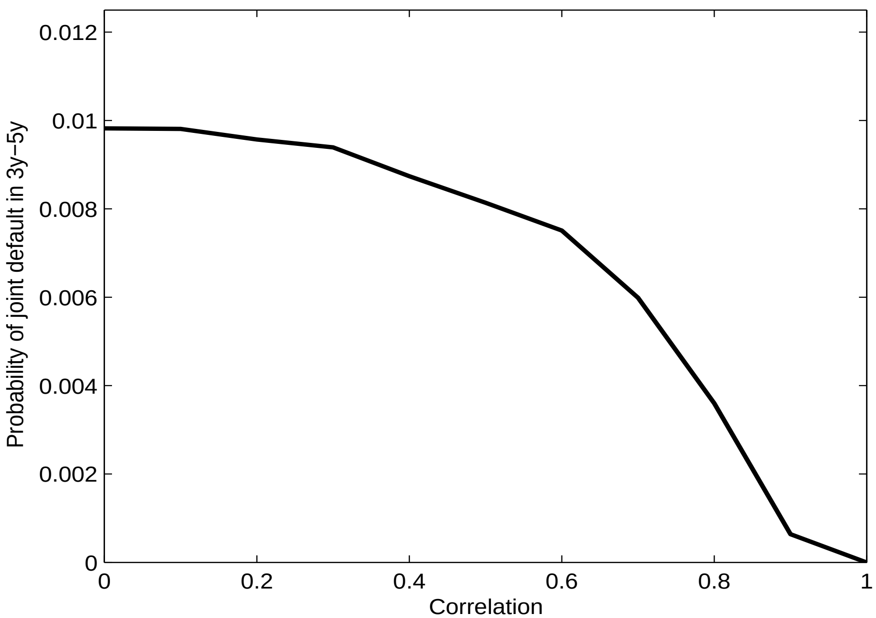
$$\lambda_1 = 0.03, \ \lambda_2 = 0.2$$

How may we protect ourselves against this risk? We may buy protection on a forward start Last-to-Default, paying us only if both names default in the specific period of time that worries us. Clearly the value of this protection depends on the probability of the two names to default together exactly in that short period of time.

$$\Pr\left(3y \leq \tau_1 \leq 5y, 3y \leq \tau_2 \leq 5y\right).$$

A paradox in market copula: missing the correlation - liquidity risk

We may expect this probability to be higher when correlation is higher, as it always happens with last to default...



The behaviour is opposite than expected. The probability of joint default in the same period of time decreases with correlation. In the standard gaussian copula approach, which is a static model, the correlation has little to do with the risk related to the timing of defaults. The supposed worst-case is actually best-case.

A paradox in market copula: missing the correlation - liquidity risk

It is not difficult to explain mathematically this behaviour: when correlation is 0, the probability that the two names default together is

$$\begin{aligned} \Pr(3 \leq \tau_1 \leq 5) \times \Pr(3 \leq \tau_2 \leq 5) &= \\ = 0.053 * 0.181 &= 0.01 \end{aligned}$$

When correlation is 1, $\boldsymbol{\tau}_1 = \frac{\varepsilon}{\lambda_1} = \frac{\varepsilon}{0.03}$ and $\boldsymbol{\tau}_2 = \frac{\varepsilon}{\lambda_2} = \frac{\varepsilon}{0.2}$, so $3 \leq \boldsymbol{\tau}_1 \leq 5$ means $0.09 \leq \varepsilon \leq 0.15$ while $3 \leq \boldsymbol{\tau}_2 \leq 5$ means $0.6 \leq \varepsilon \leq 1$, so the probability that the two events happen jointly is zero. For intermediate levels of correlation, the probability of joint default goes down monotonically from these two extremes. But this result was not immediate to guess before drawing the above chart.

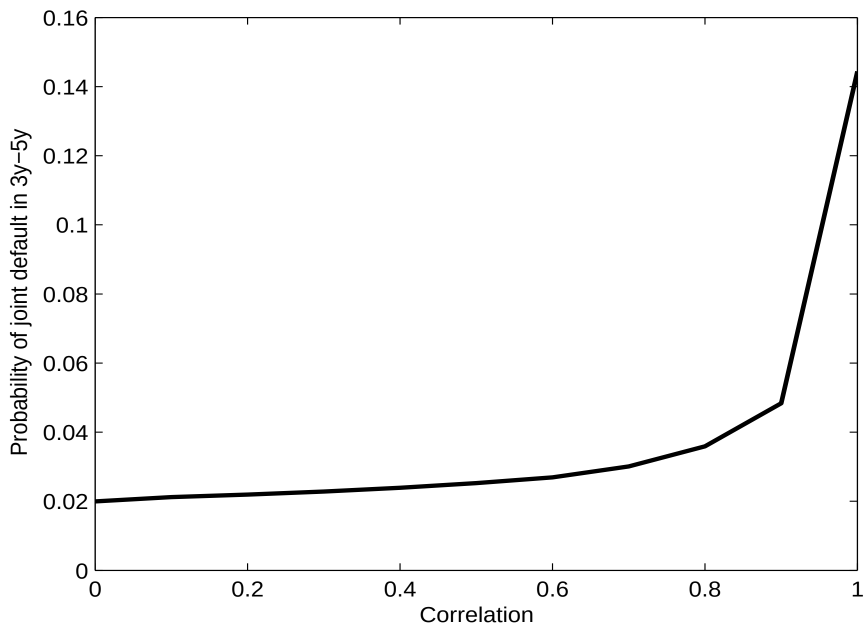
A paradox in market copula: Forward start Last to Default

For understanding better the reasons for the paradox, let suppose that, while average default risk (intensity) does not change and remains at $\frac{0.03+0.2}{2} = 0.115$, these two name to have the same risk, $\lambda_1 = \lambda_2 = 0.115$. In this case

$$\begin{aligned} \Pr(3 \leq \tau_1 \leq 5) \times \Pr(3 \leq \tau_2 \leq 5) &= \\ = 0.146 * 0.146 &= 0.021 \end{aligned}$$

while the probability that the two events happen together when $\rho = 1$ is simply $\Pr(3 \leq \tau_1 \leq 5) = 0.146$, while the pattern for intermediate values is given by the following picture.

A paradox in market copula: Forward start Last to Default



It seems that capturing the temporal aspect of "correlation risk" is related more to capturing changes in credit spreads than to gaussian copula correlation ρ.

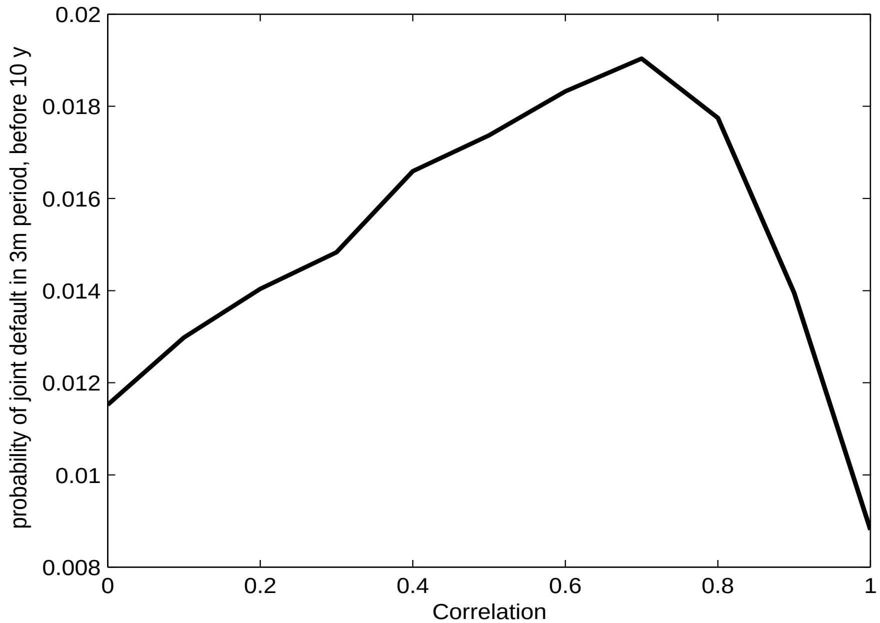
A paradox in market copula: Dynamic Var

There are other relevant calculation where this paradox can materialize. A typical Risk Management calculation is to assess the probability that some losses to exceed a given threshold in a period of time. For example, if one is exposed for the next 10 years to the default of two names, equally weighted in a portfolio, it is relevant to know not only the probability of them defaulting before 10 years, but also the probability that both defaults happen within a given number of days, say three months:

$$\begin{aligned}
&\Pr\left(Loss\,[t, t+3m] \;>\; 0.5\right) \;=\\
&\Pr\left(|\tau_1 - \tau_2| \;<\; 3m\right)
\end{aligned}$$

considering a period of 10 y, so $\tau_1 \;<\; 10y, \tau_2 \;<\; 10y$ and $t < 10y - 3m$.

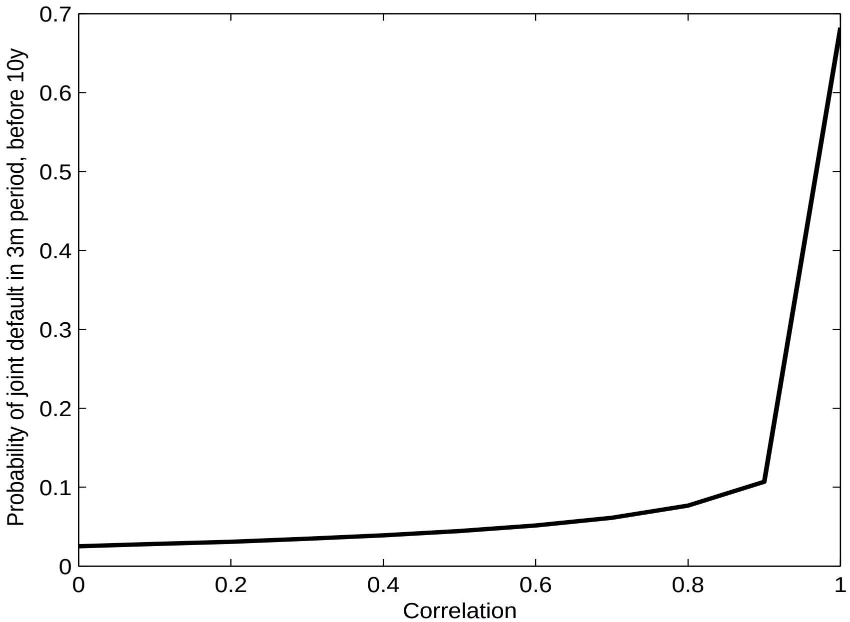
A paradox in market copula: Dynamic Var



Although this is a case different from the above forward start Last-to-default, also here we have the paradoxical behaviour that, when increasing correlation from zero to 1, at some point the probability starts decreasing when correlation increases. Here the true stress test is for correlation $\rho \approx 0.7$, not easy to guess.

A paradox in market copula: Dynamic Var

And also in this case the behaviour of correlation goes back to behave consistently with our intuition only when the two names have the same default risk.



A paradox in market copula: CDS Counterparty risk, the hard way

One more example is the assessment of counterparty risk in a CDS. The problem is tackled in Brigo and Chourdakis (2008). They concentrate on the possibility of default of a CDS counterparty before the default of the reference entity, thus focusing on the loss of the CDS mark-to-market at $\tau^{Counterparty} < \tau^{Reference}$. However there is another aspect of CDS counterparty risk which is not included in this case: it is the case

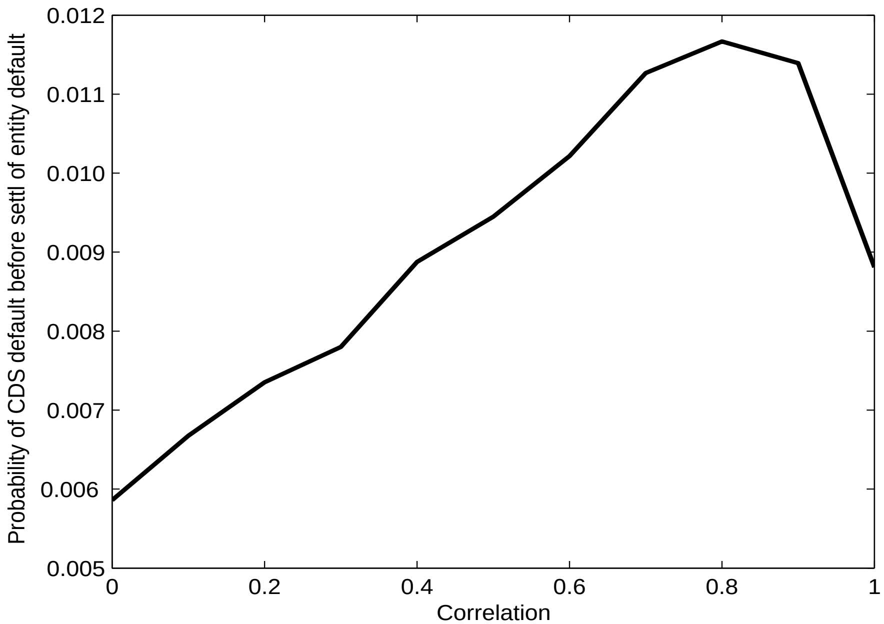
$$\begin{aligned}\tau^{Reference} &< \tau^{Counterparty},\\ \tau^{Counterparty} - \tau^{Reference} &< \Delta T^{Settlement},\end{aligned}$$

namely the case when the default of the reference entity triggers the default of the counterparty protection seller and the default of the counterparty happens before the CDS has been settled. If we consider the two names of the above example, it is natural to take name 1, with lower default risk, as the protection seller, and name 2, much more risky, as the reference entity. In this example, considering a $3m$ settlement time and a CDS with maturity 10y, the probability of suffering the heaviest CDS counterparty loss is

$$\Pr\left(|\tau_1 - \tau_2| < 3m, \tau_2 < \tau_1, \tau_1 < 10y, \tau_2 < 10y\right)$$

A paradox in market copula: CDS Counterparty risk, the hard way

Here also the probability of joint default starts decreasing when correlation approaches one, so, differently from what one may guess, $\rho = 1$ does not correspond to the riskiest scenario in terms of concentrated losses.



A paradox in market copula: CDS Counterparty risk, the hard way

- it is not easy to perform stress-test because our intuition can fail us. Only model investigation can shed some light.
- the behaviour of a model can change radically when applied to new payoff (in this case forward-starting computations).
- some models are more difficult to understand than others. A structural model where we model individual defaults and then we link them in a financially meaningful way is easier to validate and stress-test than a copula, which is a mathematical construction where individual default risks can modelled separately and then joined together with an exogenous interdependency assumption.
- The above paradoxes are even more undesirable if one considers that the temporal meaning of correlation risk, missed by the gaussian copula, has been the most relevant in the crisis. If Gaussian Copula remains the central market model for its simplicity, users should take into account the model's behaviour in the above examples.
- It is important to create a well-structured framework for stress testing and devote attention to understand what is the worst-case market scenario. However,we must also make sure that we are able to transfer it correctly into a model stress-test, otherwise results will only mislead us.

A solution: the Marshal-Olkin Model

As an alternative, we can move from the Gaussian Copula that looks like a black-box in a credit application, to a model with a clear financial meaning. An example is the model introduced by Marshall and Olkin (1967), that underlies the Marshall-Olkin Copula but does not coincide with it. Consider two names X_1, X_2 as above. For modelling the joint distribution of default times of X_1, X_2 we take three independent Poisson processes with first jumps at Z_1, Z_2, $Z_{1,2}$with parameters $\gamma_1, \gamma_2, \gamma_{1,2}$. The first jumps of each of these processes represent a financial event. The jump Z_i leads name X_i, $i = 1, 2$, to default, while a jump $Z_{1,2}$ represents an event that makes both names default. It is not difficult to write down the joint and marginal survival probabilities of this model:

$$\begin{aligned}
\Pr(\tau_1 > T_1, \tau_2 > T_2) &= \Pr(Z_1 > T_1)\Pr(Z_2 > T_2)\Pr(Z_{1,2} > \max(T_1, T_2)) \quad (20)\\
&= e^{-\gamma_1 T_1} e^{-\gamma_2 T_2} e^{-\gamma_{12}\max(T_1,T_2)}\\
\Pr(\tau_i > T_i) &= \Pr(Z_i > T_i)\Pr(Z_{1,2} > T_i)\\
&= e^{-\gamma_i T_i} e^{-\gamma_{12} T_i}.
\end{aligned}$$

A solution: the Marshal-Olkin Model

This model does not look a black-box like the Gaussian Copula. Here it is not possible to have, unlike the Gaussian Copula, events such as defaults that are perfectly dependent but the larger the first default, the further away the second. Here there are scenarios when any of the two names, independently of the other, can default for idiosyncratic reasons, and one scenario in which instead the defaults happen for a shock that affects both. In this case the two defaults happen at the same time. This also is an approximation compared to reality. But it seems more realistic than the ‘perfectly-dependent, many-years-apart default events’ of the Gaussian Copula. Here we have independence when $\gamma_{12} = 0$ and maximum dependence when $\gamma_{12} = \min(\lambda_1, \lambda_1)$, and a monotonic behaviour of the joint probability is maintained (but it is lost if we move from the model to the copula, that reintroduce the ‘perfectly-dependent, many-years-apart default events’).

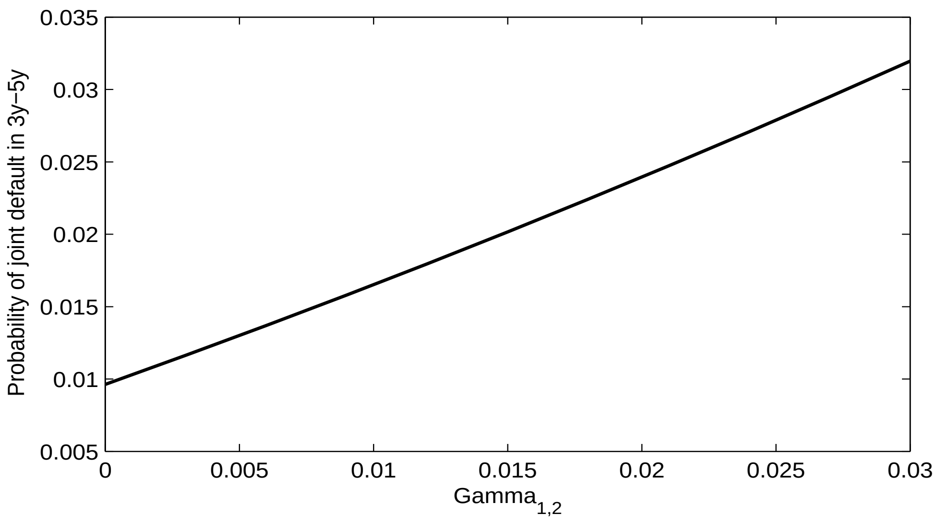


Figure 7: Probability of concentrated losses for different gamma in MO Model.

Even if the marginal intensities are different from each other $(\lambda_1 = 0.03,\ \lambda_2 = 0.2)$ and the interval has a forward start date (between 3y and 5y), in this model the parameter that controls interdependency has a clear, monotonic and stable relationship with the financial risk of interest.

Standard Mapping: missing the difference between idiosyncratic and systemic risk in skew dynamics

Here we make a "validation" of the mapping methods traditionally used for attributing a correlation to tranches of bespoke portfolios. Such methods start by using the quoted Index Tranches as a comparable, and adjust the correlations to the bespoke by assuming mathematical relations between the spreads of a portfolio and the correlation, that do not have any obvious financial meaning. There exists no liquid market for backtesting the goodness of these methods, so we devise a way to use historical information to understand if these methods capture at least the historical relation between spreads and correlation. We do not consider hundreds of dates but only a few specific dates associated to relevant market dislocation. This cannot guarantee that methods are right, but like a stress test it can tell us if they are wrong in some relevant cases. In fact the test allows to understand that standard methods see the relations between spreads and correlations only along one possible dimension (level of risk) while there has been historically another one (dispersion) that alternative methods can capture. We propose one such method to increase our benchmarks for evaluation and quantification of the model risk.

Setting Correlation for a Bespoke portfolio: Mapping?

Most portfolio credit derivatives have tranches based on *bespoke* portfolios for which there is no liquid information on implied correlations. In 2006-2007 this problem led to a characteristic market practice, that was named Mapping. In the following we describe this market technique and we devise a way of testing it on market data.

What is the idea behind Mapping? If we want to stick to Gaussian Copula with average correlation (the market model), we know that correlation is not a characteristic of the portfolio, but of the tranche. When we move to a Bespoke portfolio, we often have no liquid information on correlations, but only single name spreads. Is there a way, from Bespoke single name spreads and from Index correlation information, to understand which correlations one should give to the different tranches of the Bespoke?

According to mapping, the answer is yes.

Setting Correlation for a Bespoke portfolio: Mapping?

According to mapping, the answer is yes. A level of correlation ρ should corresponds to some fundamental characteristic of one tranche (that we call invariant), so that we can first use Index information to detect the map,

$$\rho_K^{Index} = f\left(Invariant_K^{Index}\right)$$

and then we can apply this map to the bespoke, finding bespoke correlations as functions of the invariants of the bespoke's tranches

$$\rho_{\bar{K}}^{Bespoke} = f\left(Invariant_{\bar{K}}^{Bespoke}\right)$$

What could be this invariant?

From the Index to Bespokes: Mapping Methods

1) No Mapping

In the simplest case, the invariant is just the detachment K of a tranche. This means that

$$\rho_K^{Index} = f^{NM}(K), \quad \rho_{\bar{K}}^{Bespoke} = f^{NM}(\bar{K})$$

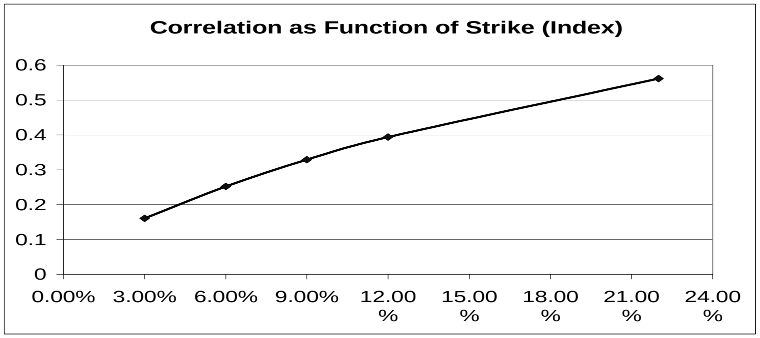
2) Expected Loss

Another solution considers not K, but $\frac{K}{\mathbb{E}[L]}$. Now one must detect from the index the map

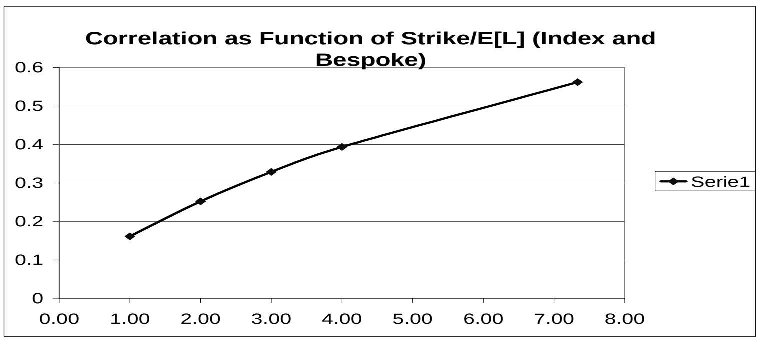
$$\rho_K^{Index} = f^{EL}\left(\frac{K}{\mathbb{E}[L]}\right)$$

and then one can apply it to the bespoke, so that

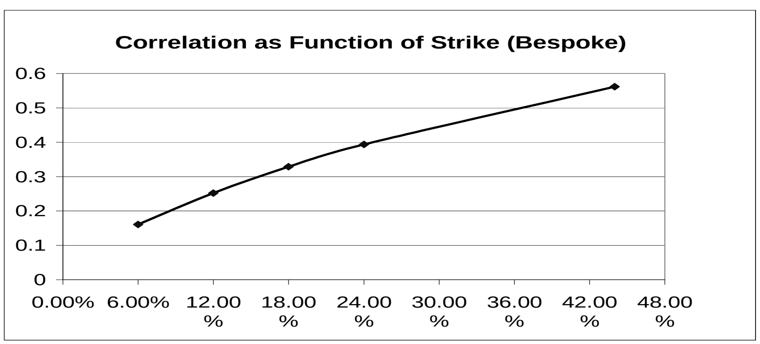
$$\rho_{\bar{K}}^{Bespoke} = f^{EL}\left(\frac{\bar{K}}{\mathbb{E}[L^{Bespoke}]}\right).$$



$E\left[L\right] = 3\%$. We rewrite the skew in terms of $K/E\left[L\right]$



This map is what DOES NOT change . We translate it back into a skew using $E\left[L^{Bespoke}\right] = 6\%$



Expected Tranched Loss

A more refined alternative uses as tranche invariant the normalized expected tranched loss $\mathbb{E}\left[L_{0,K}\right]$, where $L_{0,K}$ is the loss of a $[0\%, K\%]$ tranche

$$\frac{\mathbb{E}\left[L_{0,K}\right]}{\mathbb{E}\left[L\right]}$$

Now it is more difficult to write explicitly the map, since tranched loss itself depends on correlation. So one starts from a correlation ρ_K^{Index} in the index, and tries to find the detachment $\bar{K}$ in the bespoke such that, if we use ρ_K^{Index} for the Bespoke $\bar{K}$-tranche, we have

$$\frac{\mathbb{E}\left[L_{0,\bar{K}}^{Bespoke}; \rho_K^{Index}\right]}{\mathbb{E}\left[L^{Bespoke}\right]} = \frac{\mathbb{E}\left[L_{0,K}; \rho_K^{Index}\right]}{\mathbb{E}\left[L\right]}$$

Expected Tranched Loss

In this sense, $\bar{K}$ in the bespoke is *equivalent* to K in the Index. The equivalent strikes, through extrapolation/interpolation, provide the correlations for the bespoke.

If it is true that correlation only depends on the invariant, one can for example guess the correlation to apply to iTraxx Main or to CDX High Yield based on observing correlations on CDX Investment Grade. Let us see results of Lehman (2007), for 5y tranches, and then some tests of our own.

Lehman tests: Mapping across regions

CDX.NA.IG$\mapsto$iTraxx Main : Tranches 5y NPV

	Market	NM	EL	ETL
0-3%	10.53	10.83	9.58	10.35
3-6%	42.2	38.5	36	42.9
6-9%	12.3	7.2	6.4	10.3
9-12%	5.6	2.5	1	4.7
12-22%	2.2	0.8	0	1.9

In mapping to a portfolio of similar credit quality, although with strong regional differences, using ETL appears better than any other methods.

Lehman tests: Mapping across credit quality

CDX.NA.IG$\mapsto$CDX.NA.HY : Tranches 5y NPV

	Market	NM	EL	ETL
0-10%	68.75	61.73	74.92	74.79
10-15%	26.07	19.31	28.85	22.78
15-25%	225.7	230.2	155.2	136.7
25-35%	56.1	134.2	21.3	28.1

Instead, we see that, moving to a portfolio with a relevant difference in riskiness, mapping methods all perform badly. In particular, EL and ETL put too much risk in the equity tranche and too little in the in the senior tranche. Lehman (2007) suggest this might be related to the quotation features of the HY compared to the IG. Now we see some tests of ours, trying to clarify these issues.

Mapping: what is it, really?

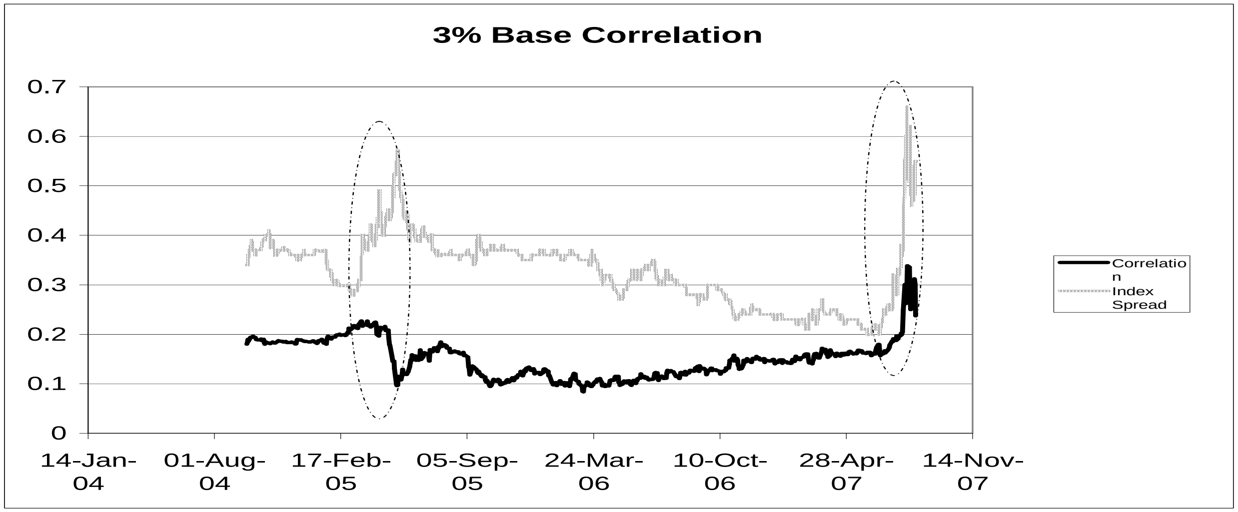
Mapping methods simply try to **capture the relation between the credit spreads of the names in a portfolio, and the base correlations that apply to the portfolio tranches**. They can work if, when moving from the Index to the bespoke, the change of the correlation skew can be captured based on the change of the spreads of the components. If mapping is applied to portfolios too different from each other, this fundamental assumption could not hold, which makes it difficult to compare the methods.

We want to test mapping for portfolios that are similar, apart from some well detected characteristics expressed by portfolio spreads. **Rather than different portfolios at the same time**, we consider **the same portfolio at different times**, in particular before and after some well detected shocks that changed both spreads and the correlation skew. We test if the mapping methods can capture at least the historical relation between spreads and the correlation skew. If a mapping method works well in this simple test, it may work also in a real Bespoke mapping. If it does not work here, it is hopeless to apply it to more complex situations.

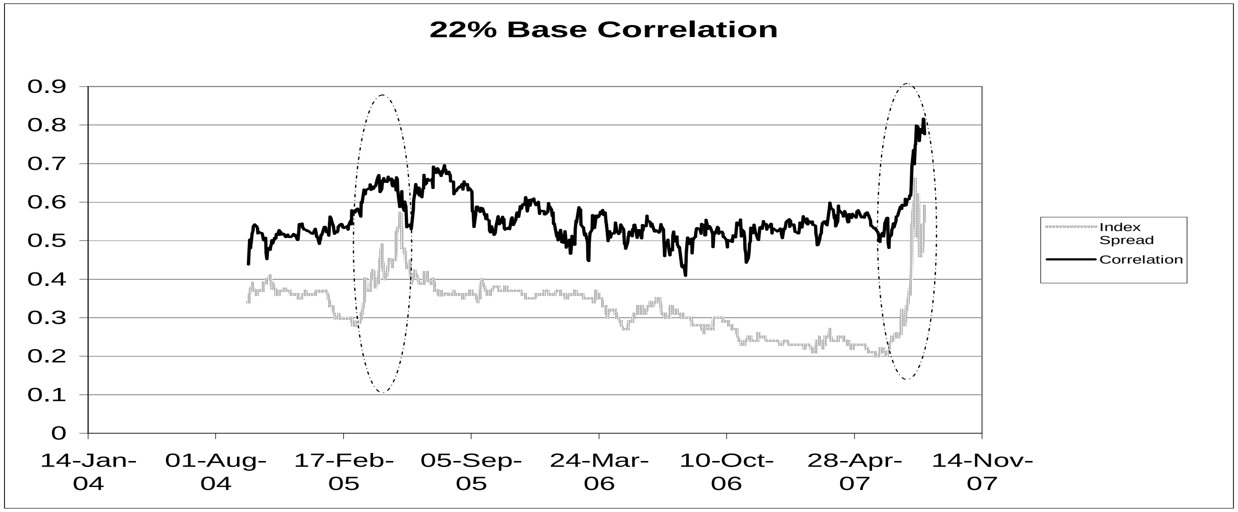
Mapping: can it also help in hedging?

If mapping allows to explain some of the historical relationship between changes in spreads and changes in correlation, this is useful also for risk-management of correlation products: if a given change in the underlying spreads tends to be associated to a specific change in correlation, it is stupid to delta-hedge (compute sensitivities wrt changes in spreads) without taking this into account. Hedging would be much more efficient if we can capture the change in correlation usually associated to change in spreads (analogous to the shadow delta in equity or IR trading).

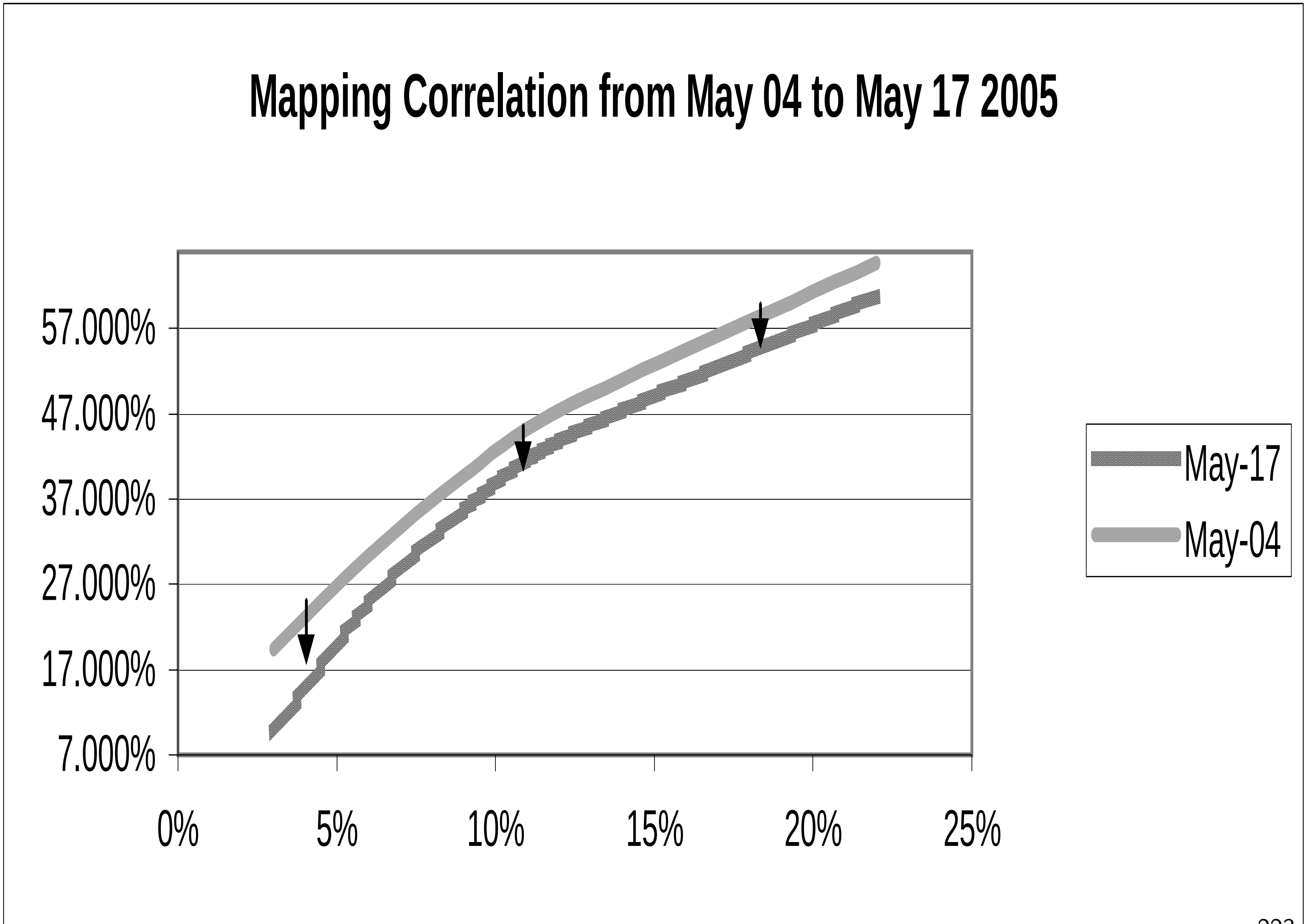
In the following we consider the iTraxx Main index to see whether the mapping methods allow to anticipate correctly (at least approximately) how the correlation skew changes when spread change. To make the job easier, we select important spread changes that were associated to relevant correlation changes.



3% Base Correlation
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
14-Jan-04
01-Aug-04
17-Feb-05
05-Sep-05
24-Mar-06
10-Oct-06
28-Apr-07
14-Nov-07
Correlation
Index Spread



22% Base Correlation
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
14-Jan-04
01-Aug-04
17-Feb-05
05-Sep-05
24-Mar-06
10-Oct-06
28-Apr-07
14-Nov-07
Index Spread
Correlation


Mapping Correlation from May 04 to May 17 2005
57.000%
47.000%
37.000%
27.000%
17.000%
7.000%
0%
5%
10%
15%
20%
25%
May-17
May-04

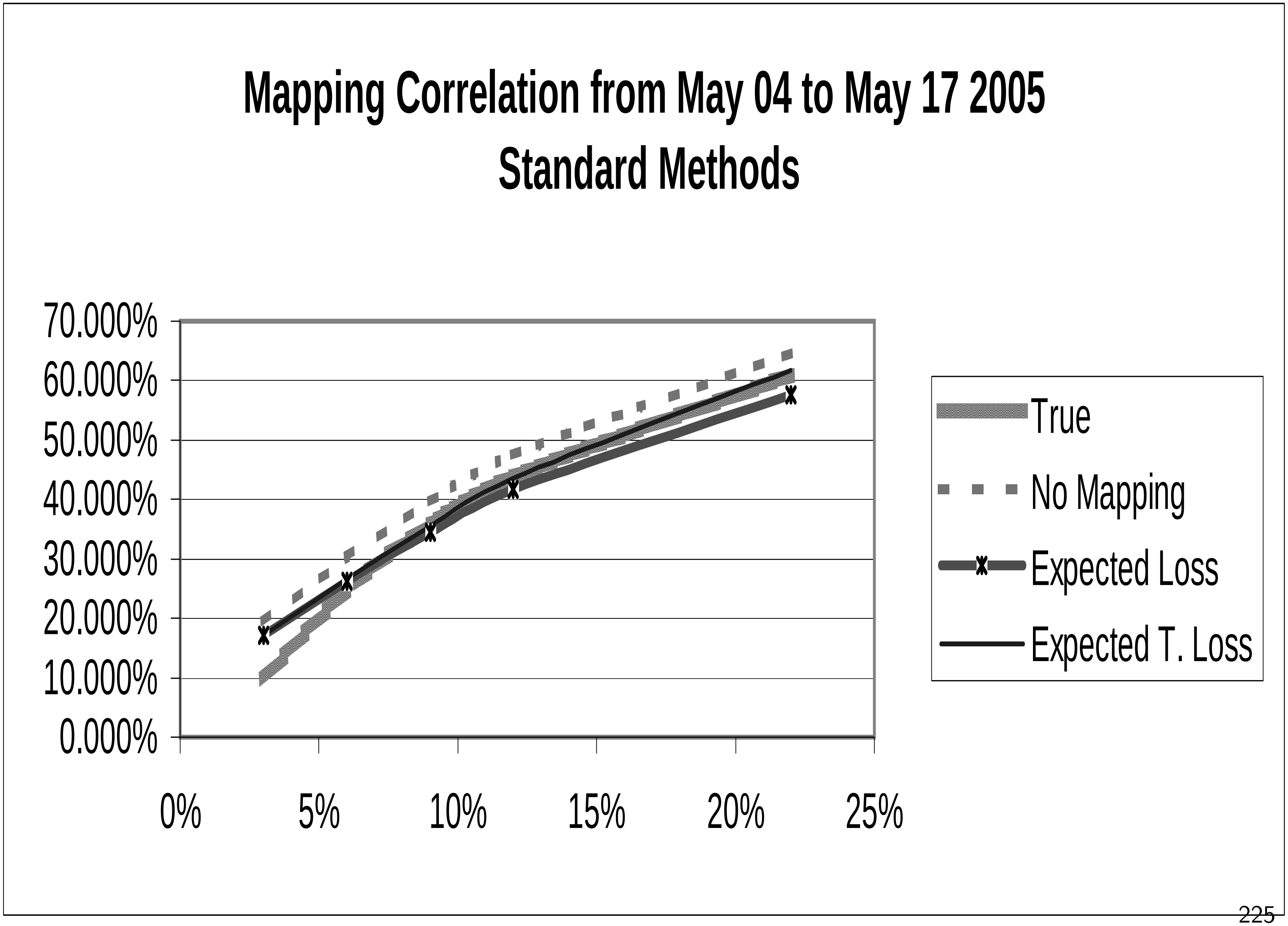
The 2005 idiosyncratic shock: Mapping when there is more idiosyncratic risk

From May 4, 2005, and May 17, 2005, the iTraxx Main average **spread increased** sharply, together with a large dislocation in the correlation skew (**correlation decreased**). It is considered a classic case of **idiosyncratic** increment of default risk.

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
Market Initial ρ	**19.38%**	**30.43%**	**39.75%**	**47.45%**	**64.72%**
Market Final ρ	**9.66%**	**24.73%**	**35.74%**	**43.78%**	**60.88%**
Difference	**-9.72%**	**-5.71%**	**-4.00%**	**-3.68%**	**-3.85%**

Let us see how one can predict it by mapping

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
No Mapping Final ρ	**19.38%**	**30.43%**	**39.75%**	**47.45%**	**64.72%**
Mapping (EL) Final ρ	**17.22%**	**26.22%**	**34.55%**	**41.72%**	**57.66%**
Mapping (ETL) Final ρ	**17.17%**	**26.62%**	**35.58%**	**43.47%**	**61.65%**

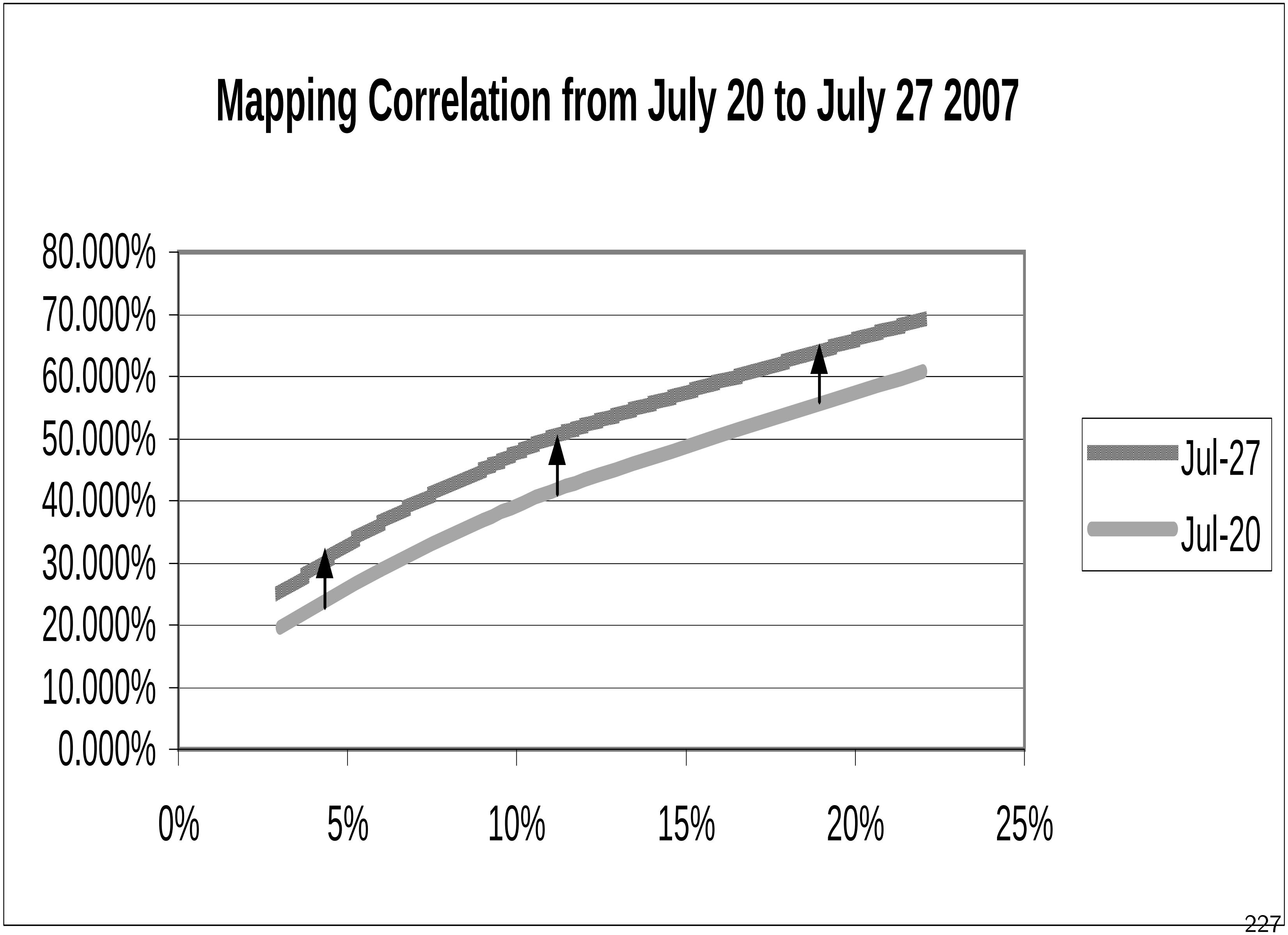

Mapping Correlation from May 04 to May 17 2005
Standard Methods
70.000%
60.000%
50.000%
40.000%
30.000%
20.000%
10.000%
0.000%
0%
5%
10%
15%
20%
25%
True
No Mapping
Expected Loss
Expected T. Loss

The 2005 idiosyncratic shock: Mapping when there is more idiosyncratic risk

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
Market Initial ρ	**19.38%**	**30.43%**	**39.75%**	**47.45%**	**64.72%**
Market Final ρ	**9.66%**	**24.73%**	**35.74%**	**43.78%**	**60.88%**
Difference	**-9.72%**	**-5.71%**	**-4.00%**	**-3.68%**	**-3.85%**

Correlation Errors	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
No Mapping	**-9.7%**	**-5.7%**	**-4.0%**	**-3.7%**	**-3.8%**
Mapping (EL)	**-7.6%**	**-1.5%**	**+1.2%**	**+2.1%**	**+3.2%**
Mapping (ETL)	**-7.5%**	**-1.9%**	**+0.2%**	**+0.3%**	**-0.8%**

Here No Mapping would induce a bigger error compared to using Mapping. Among the Mapping methods, ETL is slightly better.



Mapping Correlation from July 20 to July 27 2007
80.000%
70.000%
60.000%
50.000%
40.000%
30.000%
20.000%
10.000%
0.000%
0%
5%
10%
15%
20%
25%
Jul-27
Jul-20

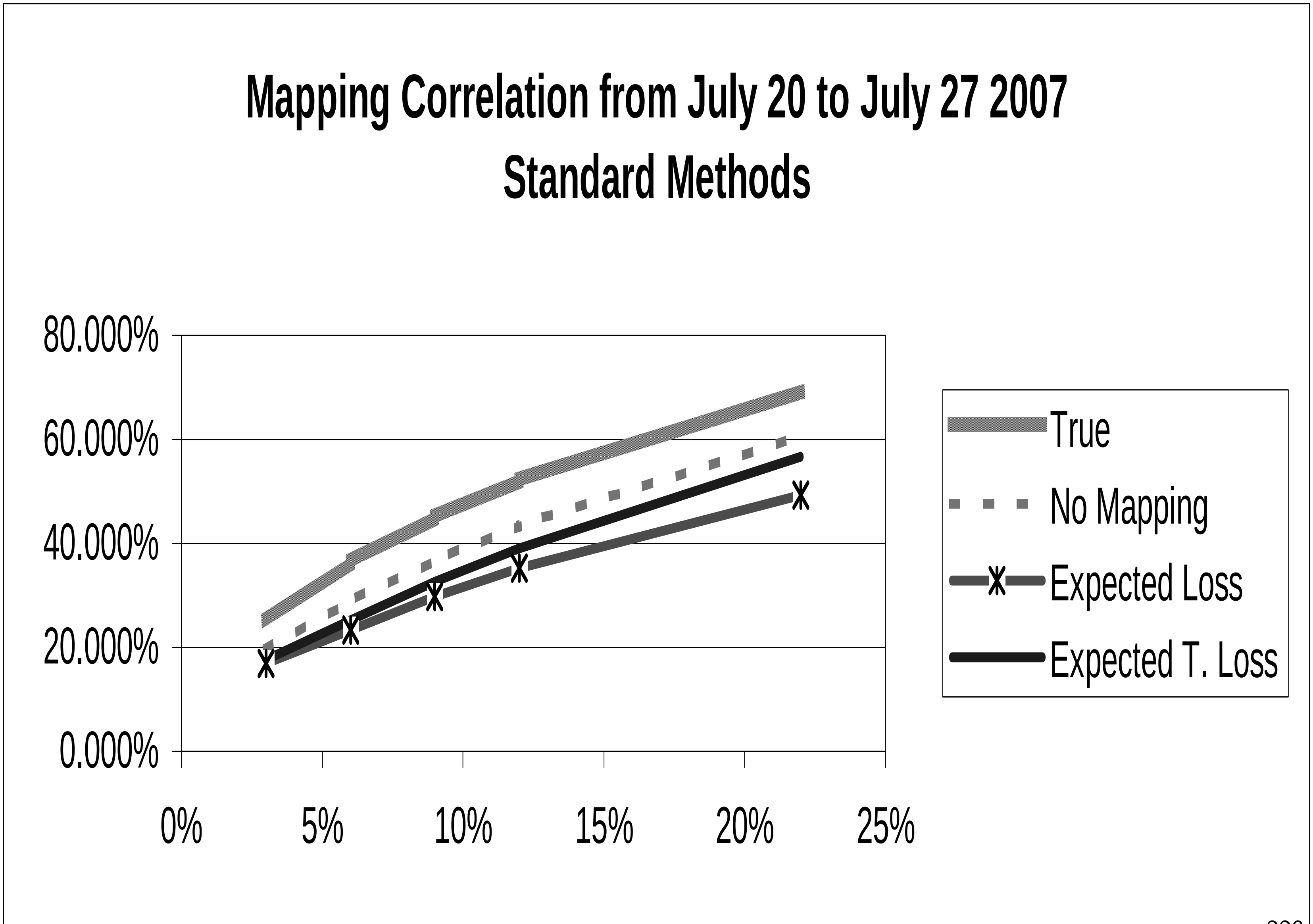
The 2007 systemic shock: Mapping when there is more systemic risk

From July 20, 2007, and July 27, 2007, we have a large increase in spreads associated to a movement in the correlation skew comparable in size to the one seen in May 2005. However this time it is an **increase of correlation**, here the increase in risk is **systemic** (it is the beginning of the credit crunch).

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
Market Initial ρ	**19.61%**	**28.93%**	**36.76%**	**43.38%**	**60.70%**
Market Final ρ	**24.92%**	**36.37%**	**44.93%**	**52.09%**	**69.34%**
Difference	**+5.31%**	**+7.44%**	**+8.17%**	**+8.71%**	**+8.64%**

Let us see how one can predict it by mapping.

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
No Mapping Final ρ	**19.61%**	**28.93%**	**36.76%**	**43.38%**	**60.70%**
Mapping (EL) Final ρ	**16.81%**	**23.33%**	**29.71%**	**35.18%**	**49.26%**
Mapping (ETL) Final ρ	**17.61%**	**25.26%**	**32.57%**	**39.11%**	**56.61%**


Mapping Correlation from July 20 to July 27 2007
Standard Methods
80.000%
60.000%
40.000%
20.000%
0.000%
0%
5%
10%
15%
20%
25%
True
No Mapping
Expected Loss
Expected T. Loss

The 2007 systemic shock: Mapping when there is more systemic risk

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
Market Initial ρ	**19.61%**	**28.93%**	**36.76%**	**43.38%**	**60.70%**
Market Final ρ	**24.92%**	**36.37%**	**44.93%**	**52.09%**	**69.34%**
Difference	**+5.31%**	**+7.44%**	**+8.17%**	**+8.71%**	**+8.64%**

Correlation Errors	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
No Mapping	**+5.31%**	**+7.44%**	**+8.17%**	**+8.71%**	**+8.64%**
Mapping (EL)	**+8.11%**	**+13.04%**	**+15.22%**	**+16.91%**	**+20.08%**
Mapping (ETL)	**+7.31%**	**+11.11%**	**+12.36%**	**+12.98%**	**+12.73%**

Here No Mapping is better than using Mapping, since the latter goes in the wrong direction, decreasing correlation (although among the mapping methods ETL is slightly better).

The 2007 systemic shock: Mapping when there is more systemic risk

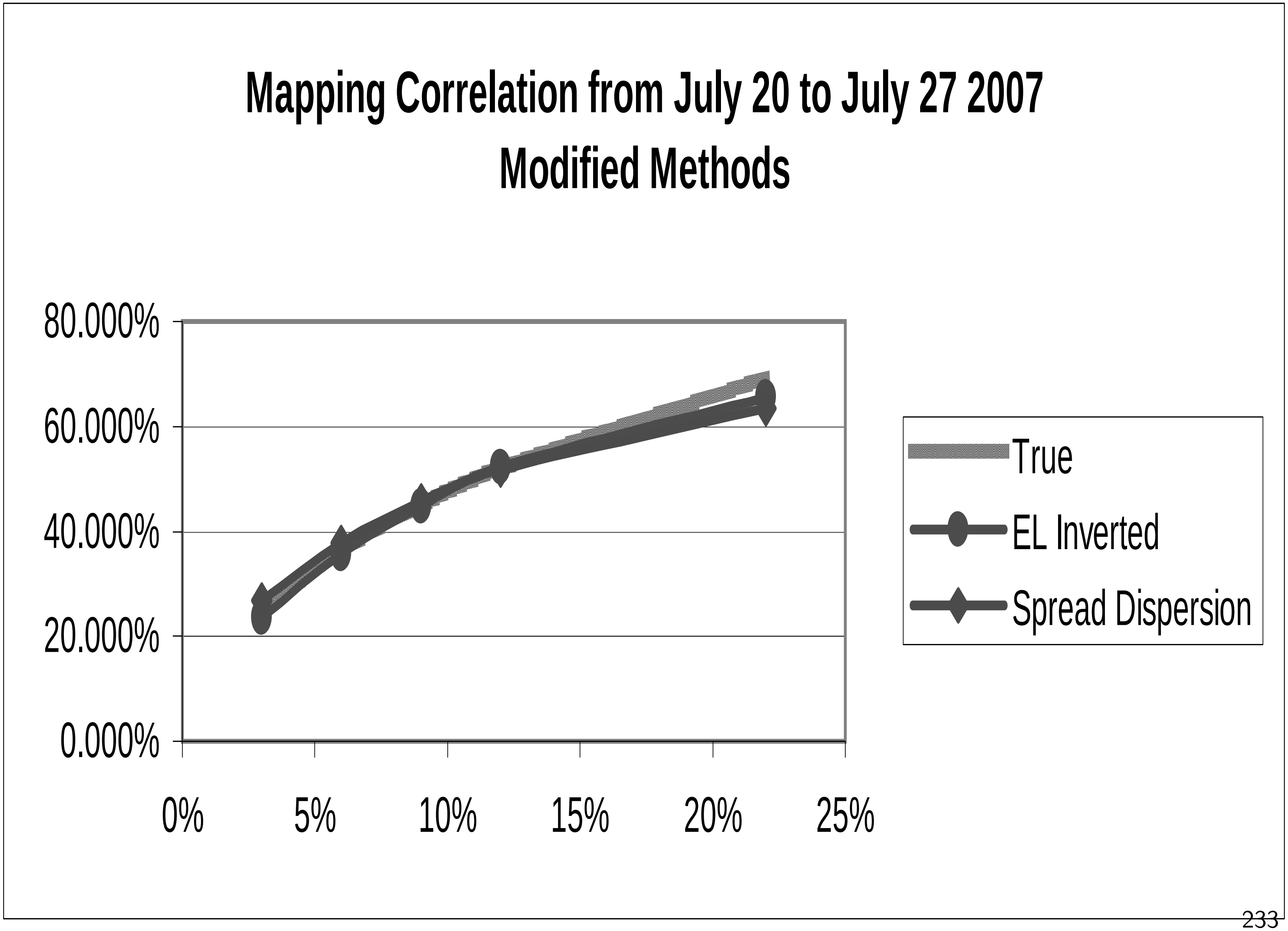
The invariant of EL Mapping ($\frac{K}{\mathbb{E}[L]}$) **can only decrease correlation when the risk of default grows**. ETL is less simplistic but has a similar behaviour. For being consistent with the actual behaviour of the market in 2007 we should radically invert the invariant of the most common mapping method. If one considers $K \times \mathbb{E}[L]$ as an invariant (Inverted Expected Loss) one has a behaviour in the right direction, but such a mapping would not work when risk of default increases because of idiosyncratic risk.

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
Market Final ρ	**24.92%**	**36.37%**	**44.93%**	**52.09%**	**69.34%**
No Mapping Final ρ	19.61%	28.93%	36.76%	43.38%	60.70%
Mapping (EL) Final ρ	16.81%	23.33%	29.71%	35.18%	49.26%
Mapping (ETL) Final ρ	17.61%	25.26%	32.57%	39.11%	56.61%
Inverted EL Final ρ	**23.61%**	**35.65%**	**44.87%**	**52.30%**	**65.46%**

The 2007 systemic shock: Mapping when there is more systemic risk

Is this the end of the mapping illusion? Standard mapping appears too give a very limited description of correlation behaviour, however it is interesting to notice that simple alternative mapping methods can give a different description consistent at least with historical behaviour. We can take as an invariant K^D, where D is the Dispersion Index computed on the portfolio spreads: $D = \frac{\mathbb{S}td(S)}{\mathbb{M}[S]\sqrt{n-1}}$ (Dispersion Method). From 20 to 27 July 2007 the dispersion index has moved from 11.28% to 8.40%. The results are

Correlations	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
Market Final ρ	**24.92%**	**36.37%**	**44.93%**	**52.09%**	**69.34%**
No Mapping Final ρ	19.61%	28.93%	36.76%	43.38%	60.70%
Mapping (EL) Final ρ	16.81%	23.33%	29.71%	35.18%	49.26%
Mapping (ETL) Final ρ	17.61%	25.26%	32.57%	39.11%	56.61%
Dispersion Final ρ	**26.76%**	**37.88%**	**45.61%**	**51.78%**	**63.36%**


Mapping Correlation from July 20 to July 27 2007
Modified Methods
80.000%
60.000%
40.000%
20.000%
0.000%
0%
5%
10%
15%
20%
25%
True
EL Inverted
Spread Dispersion

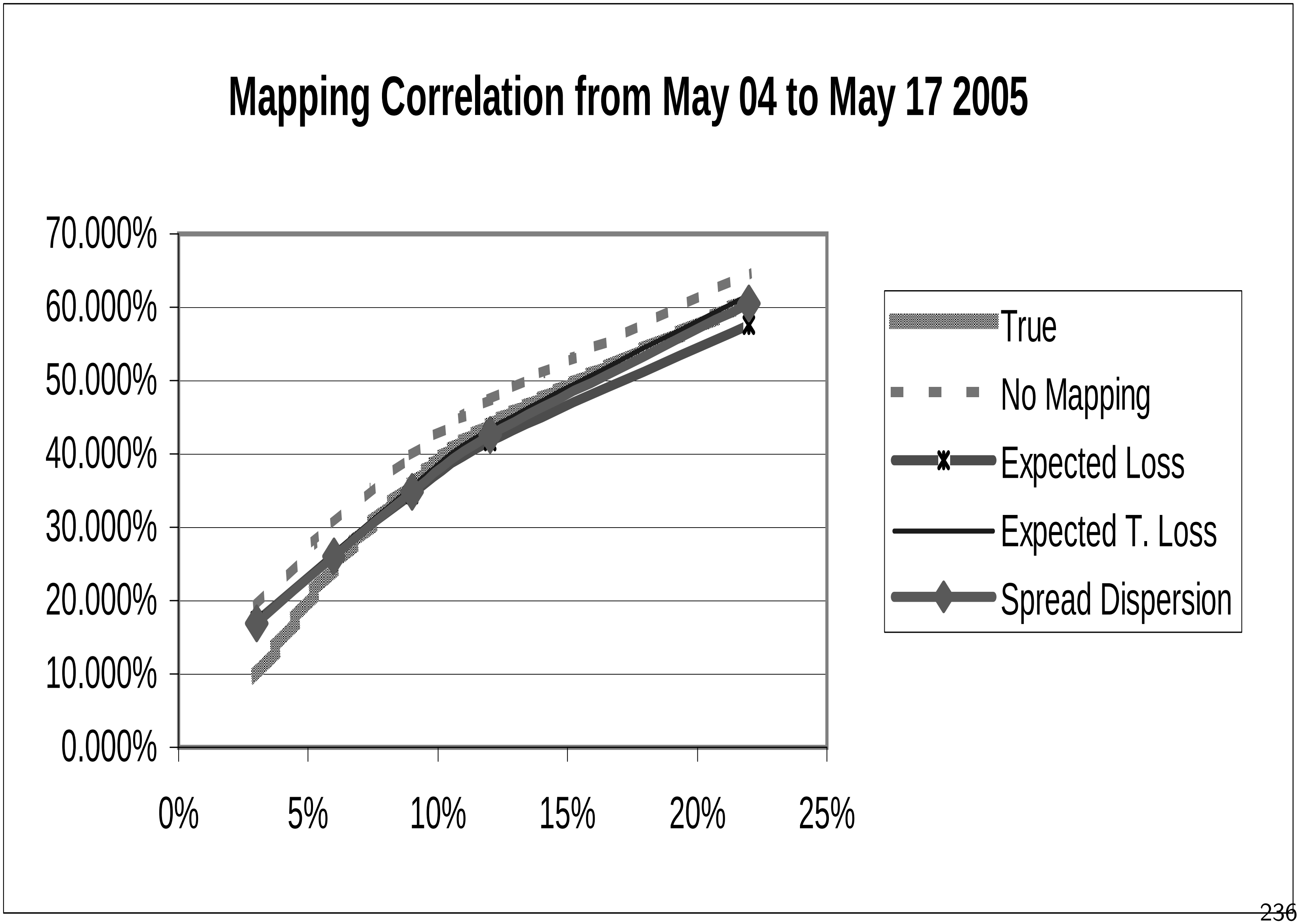
The 2007 systemic shock: Mapping when there is more systemic risk

Correlation Errors	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
No Mapping	**+5.31%**	**+7.44%**	**+8.17%**	**+8.71%**	**+8.64%**
Mapping (EL)	**+8.11%**	**+13.04%**	**+15.22%**	**+16.91%**	**+20.08%**
Mapping (ETL)	**+7.31%**	**+11.11%**	**+12.36%**	**+12.98%**	**+12.73%**
Dispersion	**-1.84%**	**-1.51%**	**-0.68%**	**+0.31%**	**+5.98%**

Intertemporal Mapping

What is interesting of the Dispersion Method is that also for the idiosyncratic 2005, when the Dispersion Index moved from 5.70% to 6.15%, it can approximately capture market regularity

Correlation Errors	*0%-3%*	*0%-6%*	*0%-9%*	*0%-12%*	*0%-22%*
No Mapping	**-9.7%**	**-5.7%**	**-4.0%**	**-3.7%**	**-3.8%**
Mapping (EL)	**-7.6%**	**-1.5%**	**+1.2%**	**+2.1%**	**+3.2%**
Mapping (ETL)	**-7.5%**	**-1.9%**	**+0.2%**	**+0.3%**	**-0.8%**
Dispersion	**-7.2%**	**-1.3%**	**+0.8%**	**+1.1%**	**+0.4%**


Mapping Correlation from May 04 to May 17 2005
70.000%
60.000%
50.000%
40.000%
30.000%
20.000%
10.000%
0.000%
0%
5%
10%
15%
20%
25%
True
No Mapping
Expected Loss
Expected T. Loss
Spread Dispersion

Intertemporal Mapping

Here the historical relation between spreads and correlation in i-Traxx is used as a kind of stress-test (rather than back-test, in spite of the use of time-series) to analyse the standard mapping methods. Standard mapping does not seem to capture the historical behaviour of spreads and correlations. In particular the standard methods appear bound to explain only idiosyncratic increases in risk. A different dmapping, more consistent with historical behaviour, can be obtained by introducing in the methods an explicit dependence on dispersion. This is useful to understand where the limits of standard methods lie and to have different methods that obey to different reasonable assumptions in order to make more aware valuations and quantification of model risk.

Model Risk in Hedging: what is hidden in a Hedging strategy

Here the focus is not on using models in valuation, but in hedging. We first observe that models that give a correct valuation can have an undesirable hedging behaviour. We consider the case of Local Volatility models in hedging options in contrast with Stochastic Volatility, based on an analysis by Pat Hagan. Then we show that

1) the actual hedging implications of a pricing model are difficult to understand, and can mislead even the model inventor,

2) the behaviour of a hedging strategy depends more on the details of the strategy than on the model used, so that the strategy requires an additional validation.

Local and Stochastic Volatility Models

In local volatility models we assume that

$$dF(t) = \bar{\sigma}(t, F(t))F(t)\ dW(t).$$

In stochastic volatility models the volatility function can remain a function of the underlying (local volatility function), but is affected by a different stochastic component,

$$\begin{aligned} dF(t) &= Vol_t\ \sigma(t, F(t))\ F(t)\ dW(t). \qquad (21) \\ dVol_t &= \mu_t dt + VolVol_t dZ(t), \quad corr\,(dW(t)dZ(t)) = \rho, \end{aligned}$$

where Vol_t is a stochastic volatility driven by its own brownian motion. Both local volatility models can price correctly options in presence of a smile. Can they also hedge correctly? The goodness of a hedging strategy depends crucially on the *dependencies that the strategy implies and their consistency with the actual market dependencies*. When a smile model is used in hedging, not even the standard delta (change in the option price for a change in F_0) is a simple quantity, because of the *shadow delta*, namely the dependency between movements of the underlying and movements of the smile.

Hedging with Local Volatility

According to Hagan et al. (2002), when the underlying F moves from F_0 to $F_0 + \varepsilon$, one expects the smile to move in the same direction (*comovement* or *sticky-delta behaviour*), as in

$$F \longrightarrow F + \varepsilon \ \Rightarrow \sigma_{F+\varepsilon}(K) \approx \sigma_F(K - \varepsilon) \tag{22}$$

so that for example

$$F \longrightarrow F + \varepsilon \ \Rightarrow \sigma_{F+\varepsilon}(F + \varepsilon) \approx \sigma_F(F)$$

Hagan at al. (2002), using singular perturbation techniques, find out that in a local volatility model with dynamics

$$dF_t = LocVol(F_t)\, F_t dW, \qquad F_0 = F$$

the implied volatility is

$$\sigma_F(K) = LocVol\left(\frac{F+K}{2}\right)\left\{1 + \frac{1}{24}\frac{LocVol''\left(\frac{F+K}{2}, ..\right)}{LocVol\left(\frac{F+K}{2}, ..\right)}(F-K)^2 + \ldots\right\}$$

The first term dominates the second one, which is usually less than 1% of the first one.

Hedging with Local Volatility

After calibration, one gets

$$\sigma_F(K) \approx LocVol\left(\frac{F+K}{2}\right)$$

With such an implied volatility function,

$$\begin{aligned} F &\rightarrow F+\varepsilon \ \Rightarrow \\ \sigma_{F+\varepsilon}(K) &\approx LocVol\left(\frac{(F+\varepsilon)+K}{2}\right) \\ &= LocVol\left(\frac{F+(\varepsilon+K)}{2}\right) \approx \sigma_F(K+\varepsilon) \end{aligned}$$

so that, for example,

$$\sigma_{F+\varepsilon}(F+\varepsilon) \approx \sigma_F(F+2\varepsilon)$$

Hedging with Local Volatility

Compare with (22). The behaviour of the model is the opposite of what desired. In particular, if the forward price F increases to $F+\varepsilon$, the implied volatility curve moves to the **left**.

Hedges calculated from the local volatility model are given by

$$\Delta^{local} = \frac{\partial \Pi\left(K, F, \sigma_F\left(K\right)\right)}{\partial F} = \frac{\partial Black}{\partial F} + \frac{\partial Black}{\partial \sigma} \boxed{\frac{\partial \sigma_F\left(K\right)}{\partial F}}.$$

THE SABR MODEL

Hagan et al. (2002) introduce the following stochastic-volatility model for solving this problem:

$$
\begin{aligned}
dF(t) &= V(t)F(t)^{\beta}\,dZ(t), \\
dV(t) &= vV(t)\,dW(t), \\
V(0) &= \alpha,
\end{aligned}
$$

where Z and W are Q^k-standard Brownian motions with $\mathbb{E}\left[dZ(t)\,dW(t)\right] = \rho\,dt$.

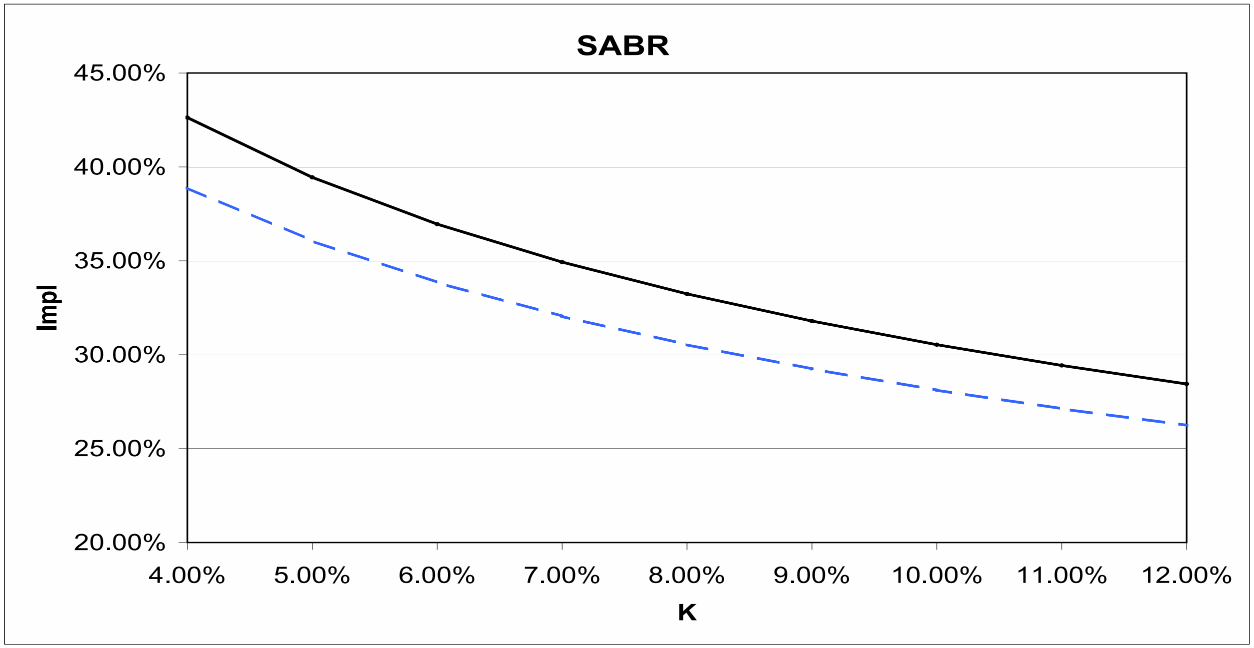
Are we sure this model implies a comovement of smile and underlying for all choices of parameters?

Hedging with Stochastic Volatility

1) Behaviour of SABR for different values of parameters:

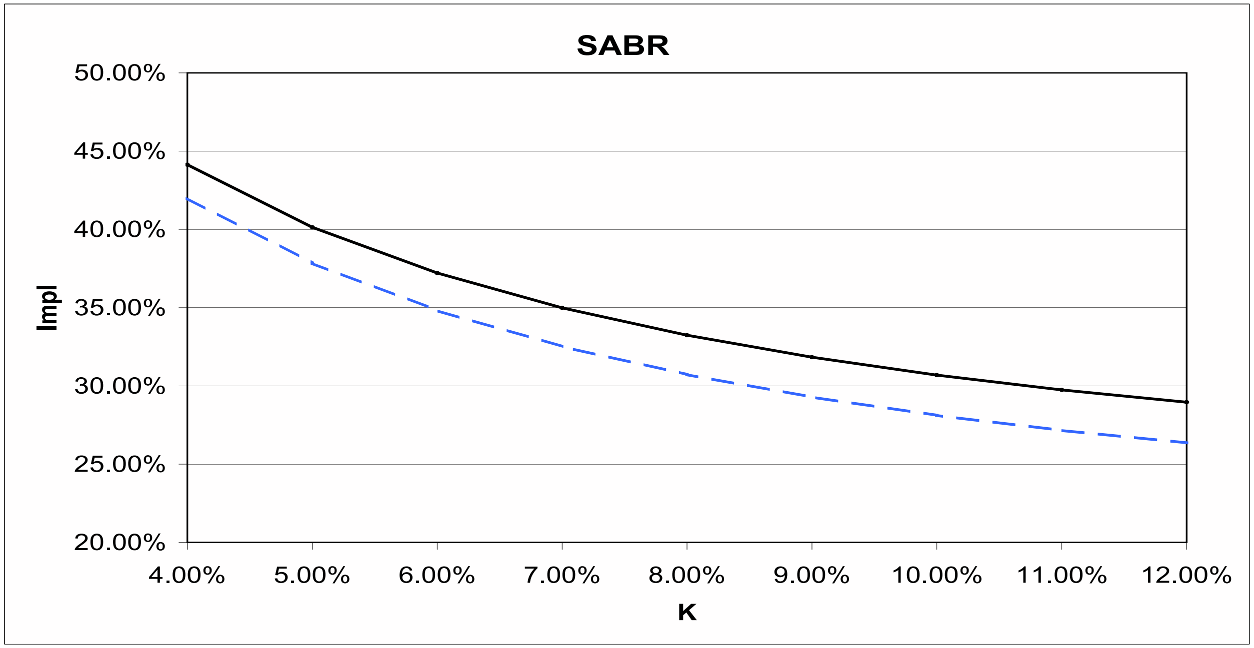
a) SABR when reduced to a local vol model (CEV):

Here $volvol = 0$, $vol = 5\%$, $\beta = 0.25, \rho = 0$. As expected, smile move backwards when $F = 8\% \rightarrow F = 10\%$.



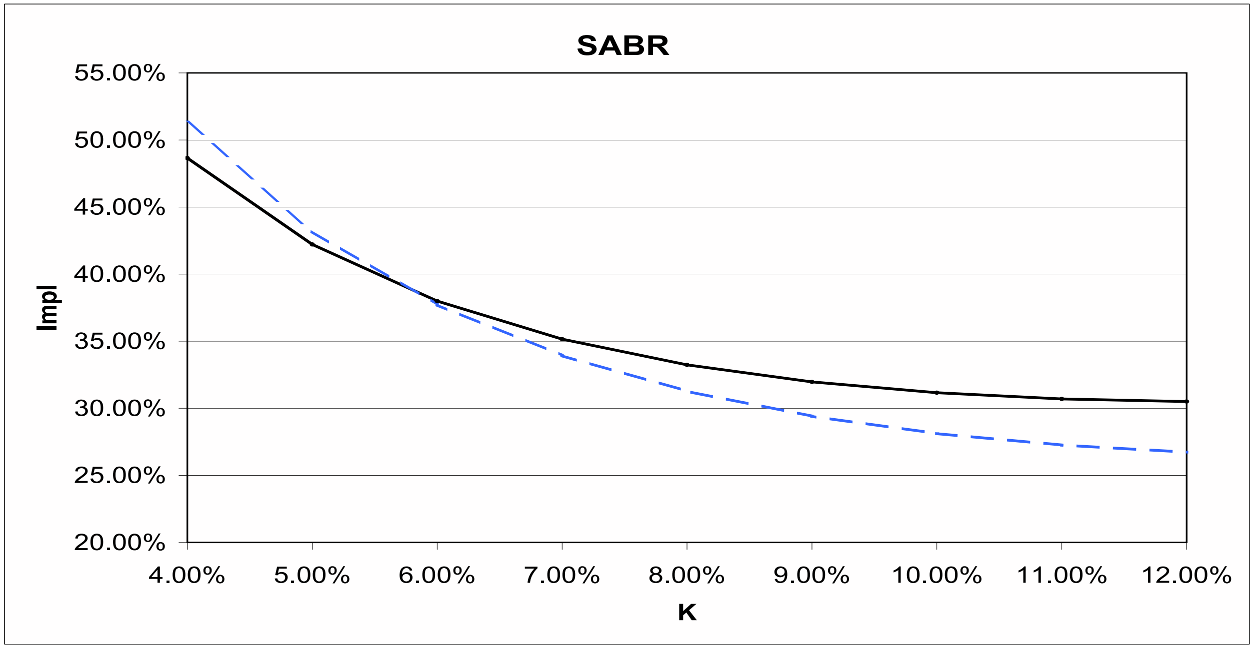
Hedging with Stochastic Volatility

a2) Here we move $volvol = 25\%$, while keeping $vol = 5\%$, $\beta = 0.25$, $\rho = 0$. This is a stochastic volatility model (with no correlation). When $F = 8\% \rightarrow F = 10\%$:



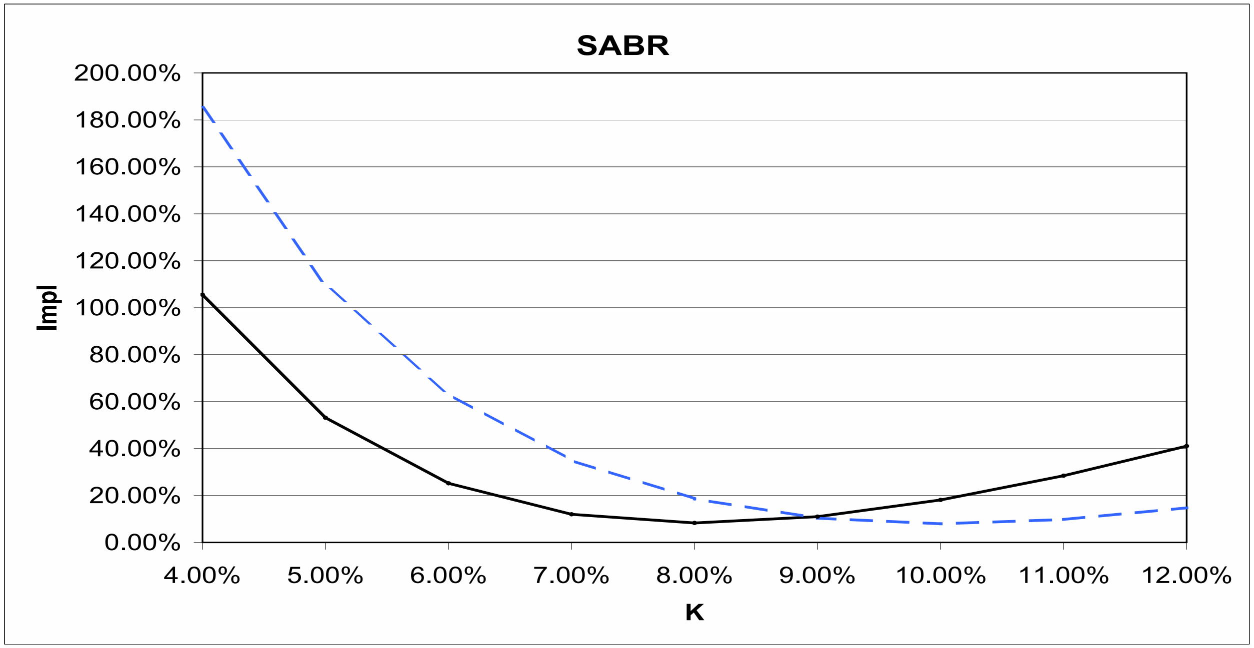
Hedging with Stochastic Volatility

b) Here we move $volvol = 50\%$, while keeping $vol = 5\%$, $\beta = 0.25, \rho = 0$. This is a stochastic volatility model with higher volvol (again no correlation). When $F = 8\% \rightarrow F = 10\%$:



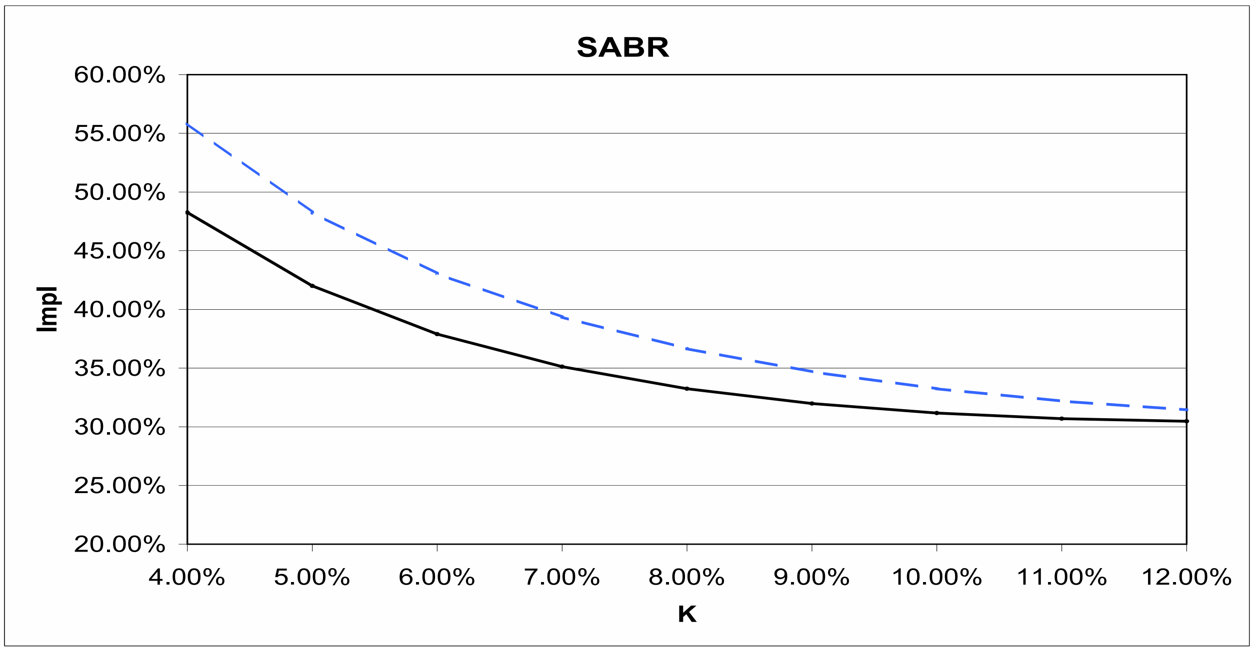
Hedging with Stochastic Volatility

c) Now we set $volvol = 100\%$ and $\beta = 0.80$, while keeping $vol = 5\%, \rho = 0$. We have moved towards lognormality, so here the stochastic volatility dominates over the skew. When $F = 8\% \rightarrow F = 10\%$:



Behaviour of SABR for different values of parameters

d) Now we change all parameters: : $volvol = 60\%$, $vol = 33.24\%$, $\beta = 1, F = 8\%, \rho = -0.41$. Here rates and stoch vol are correlated, while the local volatility is just lognormal (all skew comes from correlation). When $F = 8\% \rightarrow F = 10\%$:



Hedging with Stochastic Volatility

How to delta-hedge in a correlated stochastic volatility model? If we want to perform hedging consistently with the relations implied by the model, we have to start from the fact that a shift ΔF is due to a stochastic shock ΔW_F, and in a correlated stochastic volatility model

$$dW_V = \sqrt{1-\rho^2}dZ + \rho dW_F,$$

where $dZ \perp W_F$, so that

$$\begin{aligned} \mathbb{E}\left[dW_V \; dW_F\right] &= \rho \; dt \\ \mathbb{E}\left[dW_V \mid dW_F\right] &= \rho dW_F. \end{aligned} \tag{23}$$

Therefore assuming a shift dW_F in the underlying corresponds to expecting a contemporary shift ρdW_F in the volatility, if we are performing in-the-model hedging. How this affects the results of an hedging test? Let us see a numerical example.

We want to assess the effect of a shift of the underlying from 0.05 to 0.051. Suppose

$$\alpha = 0.1,\ volvol = 0.3,\ \beta = 1, \rho = -0.7$$

The shift in the underlying corresponds to a stochastic shock

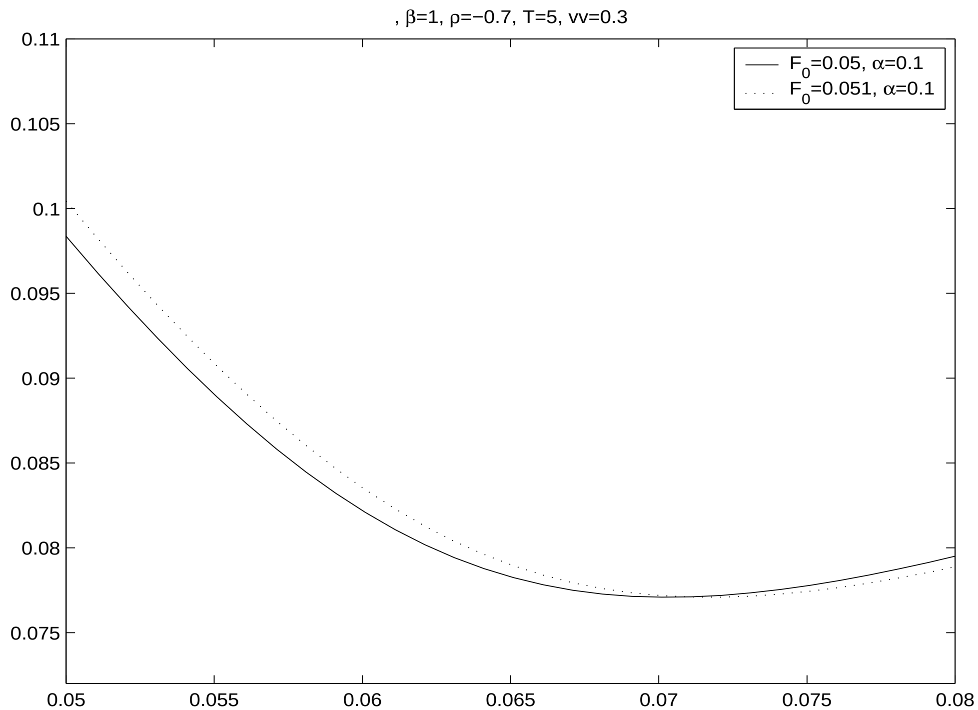
$$\Delta W_F = \frac{\Delta F}{\alpha F} = 0.2.$$

Following (23), this corresponds to an expected stochastic volatility shock

$$\mathbb{E}\left[\Delta W_V \mid \Delta W_F = 0.02\right] = \rho \Delta W_F = -0.14,$$

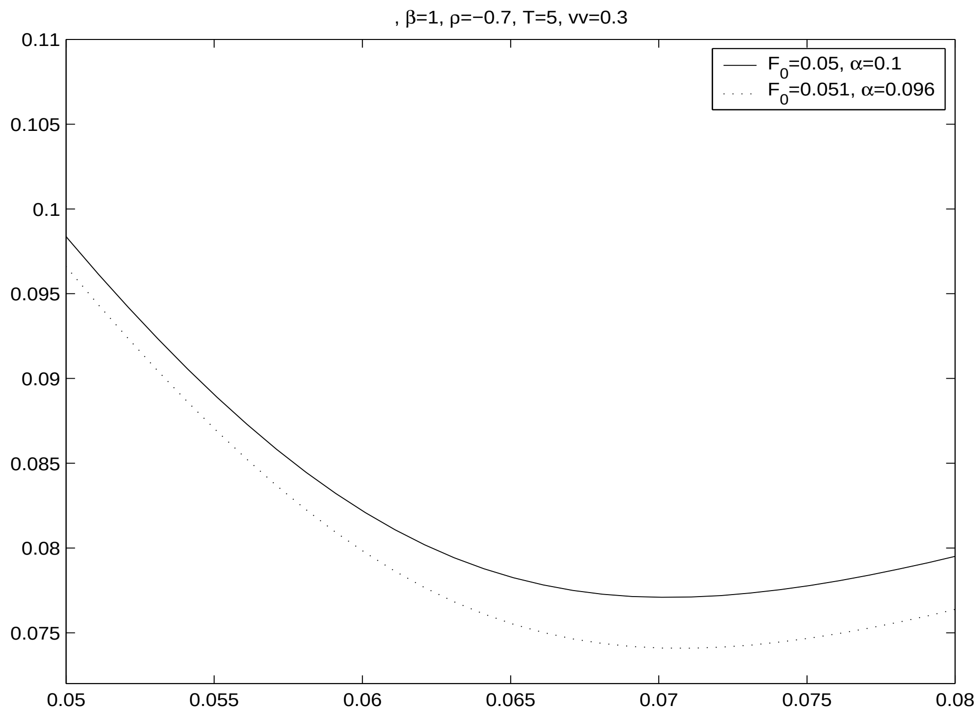
leading to a shock of the initial value of volatility α from $\alpha = 0.1$ to $\alpha \approx 0.096$.

If we neglect the expected change in volatility, as in the above hedging test mentioned above, we have the following behaviour of the model for a shift in F from 0.05 to 0.051.



This the desired behaviour: for an increase in the forward, the smile moves right.

Now we test what happen if we consider the expected change in volatility implied by the model.



In this case, the behaviour is not what we desire: the smile has moved left, and it has also decreased.

Hedging with local Volatility models

Consider the simplest possible local volatility model, the shifted lognormal,

$$dF(t) = \sigma(t)\left[F(t) + \alpha\right] dZ(t). \tag{24}$$

Marris (1999) shows that calibration to ATM options is (approximately but accurately) obtained by setting

$$\sigma = \sigma^{ATM} \frac{F(0)}{F(0) + \alpha}$$

For understanding this calibration trick, recall that this model can be derived as a dynamics in-between the Black model

$$dF(t) = \sigma^{rel} F(t)\ dZ(t)$$

and a Gaussian Model

$$dF(t) = \sigma^{abs}\ dZ(t).$$

as a simple combination

$$dF(t) = A \cdot \sigma^{rel} F(t) dZ(t) + (1 - A) \cdot \sigma^{abs} dZ(t).$$

Hedging with local Volatility models

If the Black model is calibrated to ATM options, we have simply $\sigma^{rel} = \sigma^{ATM}$. Then if we want to keep the calibration to the ATM options, we have to take

$$\sigma^{abs} = \sigma^{ATM} F(0) \tag{25}$$

so that

$$dF(t) = A \cdot \sigma^{ATM} F(t) dZ(t) + (1 - A) \cdot \sigma^{ATM} F(0) dZ(t)$$

that setting

$$\sigma = \sigma^{ATM} A \qquad \text{and} \qquad \alpha = \frac{(1 - A)}{A} F(0).$$

can be rewritten like in (24),

$$dF(t) = \sigma(t) \left[F(t) + \alpha \right] dZ(t).$$

This shows why calibration to ATM options corresponds to $\sigma = \sigma^{ATM} A = \sigma^{ATM} \frac{F(0)}{F(0)+\alpha}$.

Hedging with local Volatility models

$$dF(t) = \sigma(t)\left[F(t) + \alpha\right] dZ(t), \qquad F(0) = F_0$$

Consistency requires that I hedge simply by changing the initial condition from $F(0) = F_0$ to $F(0) = F_0 + \varepsilon$, leaving all the rest unchanged, in particular α which is just a model parameter.

However a frequent market standard recalls the fact that calibration of the model to ATM options requires

$$\alpha = \frac{(1 - A)}{A} F(0)$$

so as to write

$$dF(t) = \sigma(t)\left[F(t) + \alpha\left(F(0)\right)\right] dZ(t), \qquad F(0) = F_0$$

and hedge by changing the initial condition from $F(0) = F_0$ to $F(0) = F_0 + \varepsilon$ and also by changing from $\alpha\left(F(0)\right)$ to $\alpha\left(F(0) + \varepsilon\right)$. This is not a model consistent hedging (it hides a recalibration).

Hedging with local Volatility models

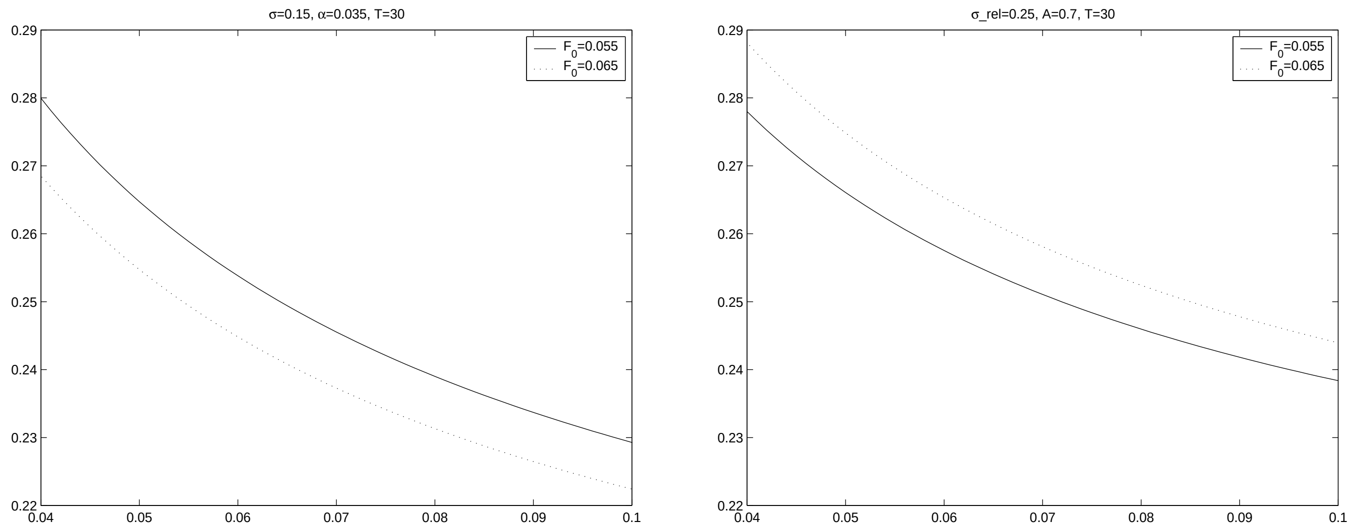


Figure 8: The shifted lognormal: In and Out of the model Hedging

Hedging with local Volatility models

The analysis shows that the desirable behaviour of SABR in hedging was not a model feature but the effect of an inconsistent hedging, ignoring relations that are in model assumptions. Both local and stochastic volatility model behave in a wrong way when used consistently, but both are usually applied inconsistently to hedging obtaining naturally a desirable behaviour. Should we really care about model consistency in hedging? Traders know that a model will be recalibrated tomorrow, therefore they do not care too much about being consistent with model assumptions in computing sensitivities. On the other hand, they build sensitivites that try to anticipate the effect of tomorrow's recalibration.

"*Pricing models can be reliable for representing changes in derivatives prices given changes in more fundamental market variables, but not reliable as a representation of the dynamics of market variables themselves*" (Trader's view)

"*Most derivatives dealers tend to believe there are too few factors in a model to sufficiently capture the evolution of the underlying*" Li (1999)

However, if a model is applied to hedging assuming relations different from those used in pricing, the appropriateness of a model for valuation has little to do with its hedging efficiency, and the two things require separate analysis.

Explaining Basis Spreads and what has happened in the Interest Rate Market

The crisis and interest rate derivatives

At the beginning of the credit crunch a number of relationships between market quotes that had been there for at least a decade suddenly broke down. This created serious turmoil, since such relationships were the foundation of hedging strategies and of the construction of the yield curve used by every bank to price derivatives of all asset classes. Let us see in more detail the relationships that broke down at the beginning of the credit crunch.

Forward contracts and Basis Swaps

Define $L(t,T)$ the Libor rate set at t for maturity T. In this work we consider Forward Rate Agreements (FRA). A FRA with fixing T, payment time T' and fixed rate K has a payoff at T' given by

$$\left[L\left(T,T'\right)-K\right]\left(T'-T\right).$$

We indicate by

$$F\left(t;T,T'\right)$$

the equilibrium level of K quoted in the market at time t.

We consider also Basis Swaps (BS). A Basis swap is a contract where two parties exchange two floating legs, both with last payment at T, based on two different payment tenors, α and α', $\alpha<\alpha'$. The α leg has a higher frequency of payment, and pays Libor rates with a shorter tenor. A spread Z is added to the payments of the α leg.

We indicate by

$$B\left(t;\alpha/\alpha';T\right)$$

the equilibrium level of Z quoted in the market at time t.

The problem

Our analysis starts from observing that some long-standing relationships regarding these quotes broke down at the burst of the credit crisis. Define the following functions of the market Libor rate:

$$
\begin{aligned}
P_L(t,T) &= \frac{1}{1+L(t,T)(T-t)}, \\
F_L\left(t;T,T'\right) &= \frac{1}{T'-T}\left(\frac{P_L(t,T)}{P_L(t,T')}-1\right) \\
&= \frac{1}{T'-T}\left(\frac{1+L\left(t,T'\right)\left(T'-t\right)}{1+L(t,T)(T-t)}-1\right)
\end{aligned}
$$

The problem

Before the crisis, the market evidence was

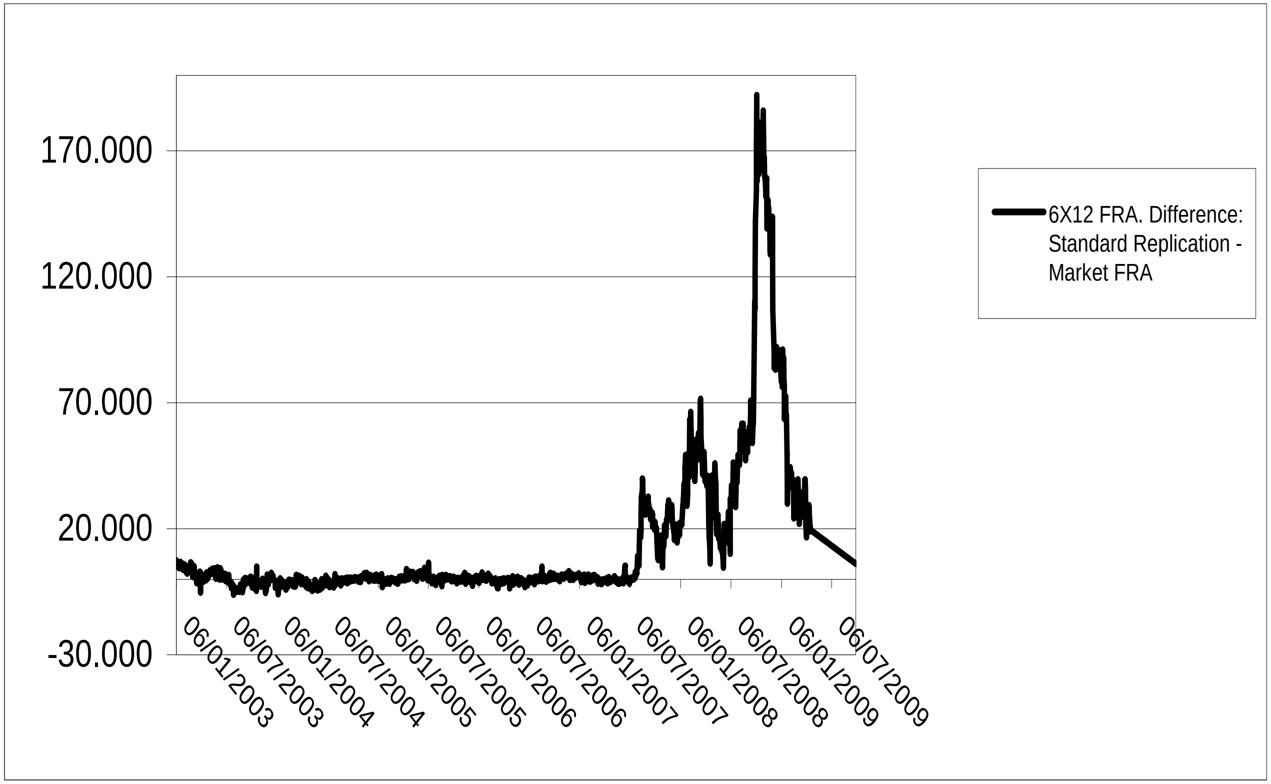
$$\begin{aligned} F\left(t;T,T'\right) &\approx F_L\left(t;T,T'\right), \\ B\left(t;\alpha/\alpha';T\right) &\approx 0 \end{aligned} \tag{26}$$

and this was explained in textbooks by replication arguments. After the crisis (from July 2007), the market evidence became

$$\begin{aligned} F\left(t;T,T'\right) &\ll F_L\left(t;T,T'\right), \\ B\left(t;\alpha/\alpha';T\right) &\gg 0 \end{aligned} \tag{27}$$

August 2007: the replication breaks down

Below we report the difference between $F_L(t; t+6m, t+12m)$ and $F(t; t+6m, t+12m)$ with t covering a period of more than 6 years.



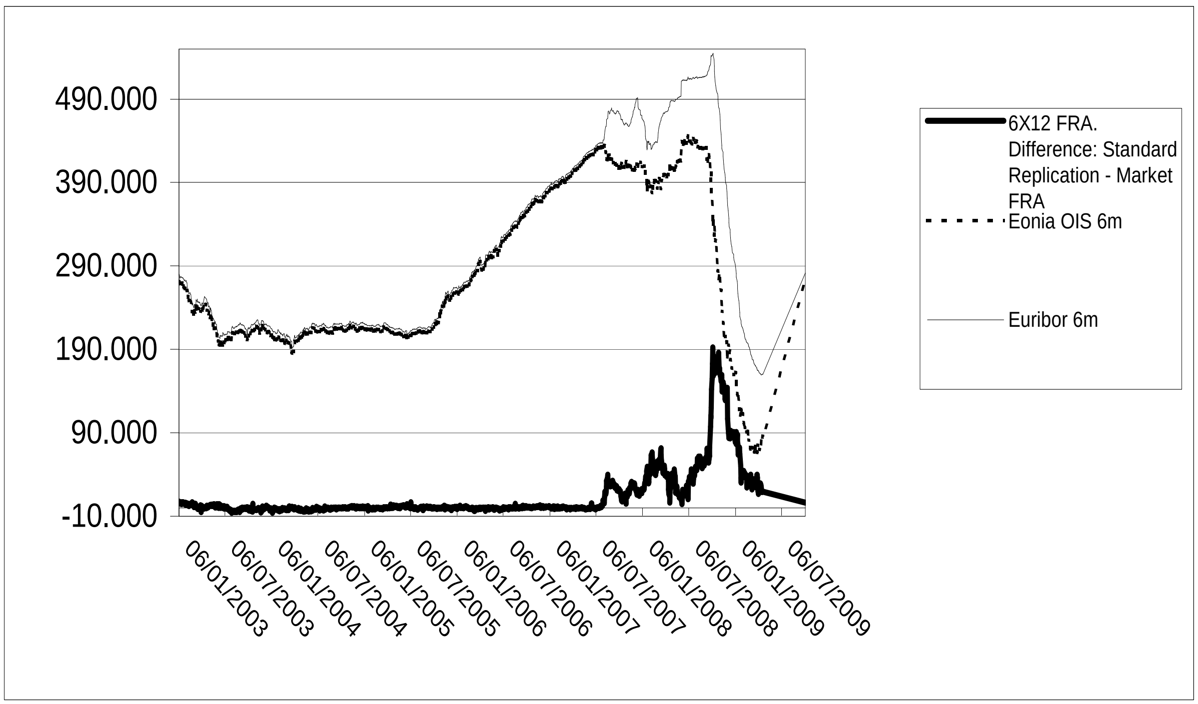
August 2007: the replication breaks down

The market FRA rate F and the Standard Replication F_L never exactly coincide, but the difference averages 0.88bp (0.000088) in the three years preceding July 2007. After July 2007, a gap $F_L - F$ opens, and remains clearly positive, averaging to 50bp from August 2007 to May 2009.

The pre-crisis relationships (26) are consistent with the results that one obtains assuming that the market is free of default and liquidity risk (the risk-free model of the market typical of standard interest rate derivatives modelling). The breakdown of the relations at the beginning of the crisis suggests that this market model is not valid anymore. Therefore we will construct alternative models of the market to see if they can explain why relations (26) do not hold anymore and why they have been replaced by (27). In particular we will introduce default and liquidity risk in the model of the market.

The wedge between Libor and OIS

We focus on these two risk factors since the discrepancies (27) erupted when another major discrepancy arose in the market: the one between Libor (Euribor) and Eonia OIS (Overnight Indexed Swaps) rates.



OIS rates as risk-free rates?

An OIS is a fixed/floating interest rate swap with the floating leg tied to a published index of a daily overnight reference rate, for example the Eonia rate in the Euro market. According to experienced market operators, an overnight rate incorporates negligible credit or liquidity risk. An OIS rate with maturity T can be seen as an average of the market expectations about future overnight rates until T, so it somewhat 'extends' overnight rates to longer terms remaining free of credit or liquidity risk. Since Libor is instead considered a rate that incorporates some credit/liquidity risk in the interbank market, the market interpreted the explosion of the Libor-OIS gap as a sign of the explosion of these two risks in the rates market.

Libor as a risky rate?

You may find the above statements rather generic. Why should an overnight rate be free of default or liquidity risk? And about libor, what does it mean that it incorporate some credit/liquidity risk? Libor does not refer to any specific counterparty; does it depend on some average liquidity/credit risk? What are the companies to which this average refers?

Many claims we make about the interest rate market are actually imprecise, when we start introducing credit and liquidity. In fact we have interest rate theory for a perfect market with no risks such as default. In a such a market, as we will see, there exists naturally one single interest rate even if there are thousands of counterparties. Here we try to go beyond this situation by introducing different models of the interest rate market.

Credit or Liquidity?

The one thing where we simplify is the distinction between credit and liquidity. It is difficult to disentangle liquidity risk from credit risk, in particular when one is analyzing not the not a single derivative deal but the interest rate market as whole. Liquidity risk can be either

1. funding liquidity risk (the risk of running short of available funds). Funding liquidity risk for a bank is normally strongly correlated to its own risk of default, since an increase of the cost of funding of a bank is usually both a cause and consequence of an increase in risk of default
2. market liquidity risk (he risk of having large exposures to markets where it is difficult to sell a security). The interbank interest rate market that is our underlying is the market of deposits, cash loans from a bank to another one. Therefore market liquidity is here the difficulty to transfer a cash loan with a specific counterparty, and as such it is always strongly correlated to the risk of default of the counterparty.

In the first part we do not attempt to separate precisely liquidity from credit in the spreads over risk-free rates. In the end, however, we outline how the two components may be separated and we study better their relationship.

Libor as an index rate

There is one first thing we have to introduce to make our models minimally realistic. Interest rate derivatives never have one single underlying, unlike what we usually pretend in our model. Even if we consider the above FRA, where the underlying is the rate $L\left(T, T'\right)$, the product is in reality more similar to a basket or index derivative. In fact the rate $L\left(T, T'\right)$, a Libor rate in this example (other derivatives may have as an underlying other rates such as Euribor which are all pretty similar to Libor), comes out of a rather elaborate process.

The Libor Market

The level of Libor is provided by the fixings, *trimmed* averages of contributions from a panel including the most relevant banks in the market *with the highest credit quality*. The banks contribute the rate at which they can borrow money in the market. The borrowing must be *unsecured*, therefore it corresponds to a *deposit* contract. Usually the quotes in the deposit market track closely the Libor fixings, since banks should simply contribute the funding rate they see in the deposit market, and the players in the deposit market are normally a set larger than Libor contributors but similar in composition. Indeed, since deposits are unsecured only borrowers with high credit quality can access this market. This is why the deposit market is also called the Libor market. In the following, when we speak of Libor banks we mean indifferently banks belonging to the panel or only to the Libor market.

The way we interpret the Libor rates in our models is crucial to the pricing of even the simplest interest rate derivative, the swap. This issue has never been deeply investigated since it is trivially solved when the market is risk-free, or even when there is a homogeneous and stable amount of credit risk in interbank transactions, as we will see. The issue becomes non-trivial when neither of these conditions hold.

Which model of the market explains the new patterns?

In the following we are going to consider different models of the interest rate market. Each model will bring about different relations between different objects, that we will compare with the actual relations we have in the market to understand which model is a better representation of reality.

We will try to use a notation that recalls the assumptions that underly a given relations. For rates and bonds we indicate if they are considered risk-free (because the counterparty cannot default or because the deal is collateralized):

$$L^{rf}(t,T), P^{rf}(t,T),$$

or if they are defaultable:

$$L^{dA}(t,T), P^{dA}(t,T),$$

where d stands for "defaultable" and A indicates the counterparty. In fact, when there is default risk one needs to specify the counterparty.

Which model of the market explains the new patterns?

We have also derivative quantities, like the equilibrium rate of a forward rate agreement or the equilibrium basis spread $B_?^? \left(t; \alpha/\alpha'; T\right)$. These quantities are different depending on:

- the fact that the product is subject to counterparty risk or not, recalled by the superscript

$$F_?^{rf} \left(t; T, T'\right), B_?^{rf} \left(t; \alpha/\alpha'; T\right) \text{ or } F_?^{dA} \left(t; T, T'\right), B_?^{dA} \left(t; \alpha/\alpha'; T\right)$$

- the assumptions made on the underlying Libor quote, recalled by the subscript. We consider three main cases:

1. Libor market free of default risk: $F_{rf}^? \left(t; T, T'\right), B_{rf}^? \left(t; \alpha/\alpha'; T\right)$
2. Libor market with stable default risk $F_{dX_0}^? \left(t; T, T'\right), B_{dX_0}^? \left(t; \alpha/\alpha'; T\right)$, where Libor is linked to the default risk of a counterparty X_0 chosen at 0.

3 Libor market with volatile default risk

$$F^{?}_{dX_0,X_\alpha,\ldots}\left(t;T,T'\right),\, B^{?}_{dX_0,X_\alpha,\ldots}\left(t;\alpha/\alpha';T\right)$$

where Libor is linked to the default risk of a sequence of counterparties.

The Riskless Interest Rate market

Let us start from the classic assumption of absence of default risk. We call $L^{rf}(t,T)$ the fair simply-compounded interest rate applying to a loan from a bank to another bank from t to T when the loan is free of default and liquidity risk. In this case the rate $L^{rf}(t,T)$ gives an indication of the time-value of money, namely what, according to the market consensus, is the compensation that a lender must ask for giving up his money for the period from t to T. If the market is arbitrage-free, the same rate $L^{rf}(t,T)$ applies whichever bank we choose as a borrower, because there are no bank-specific issues, like default risk, that can justify different quotes for the same payoff.

The Riskless Interest Rate market

If we assume that the entire interbank lending market is free of default risk, the meaning of the Libor rate $L(t,T)$ is clear. All banks borrow at the same rate $L^{rf}(t,T)$ that embeds the time-value of money, therefore

$$L(t,T) = L^{rf}(t,T) \tag{28}$$

The risk-free bond has a price

$$P^{rf}(0,T) = \mathbb{E}\left[D(0,T)\right],$$

In an arbitrage-free market the rate $L^{rf}(0,T)$ that applies to a loan for a notional equal to 1, needs to satisfy

$$1 = \mathbb{E}\left[D(0,T)\left(1 + L^{rf}(0,T)\,T\right)\right] = P^{rf}(0,T)\left(1 + L^{rf}(0,T)\,T\right),$$

so that

$$P^{rf}(0,T) = \frac{1}{1 + L(t,T)(T-t)} = P_L(t,T).$$

Here fixings and Deposit quotes simply coincide, being two ways of giving information on $L^{rf}(t,T)$.

Pricing a swap by replication

This leads to the possibility of very simple replication procedures to price swaps. A standard spot-starting swap with first fixing today at $t = 0$, tenor α (say 6 months) and M payments involves payments at the set of dates

$$\alpha, 2\alpha, \ldots, M\alpha.$$

The fixed leg is trivially evaluated. Neither is it difficult to evaluate a floating leg, since it is very simply replicated by basic instruments.

Consider a notional equal to 1. An investor can
1) borrow today an amount $P^{rf}(0, M\alpha)$ agreeing to give back 1 at maturity $M\alpha$
2) lend 1 to an interbank counterparty until α.
At α the investor receives $1 + L^{rf}(0, \alpha) = 1 + L(0, \alpha)$, and lends 1 again to the counterparty until 2α. At 2α the investor receives $1 + L^{rf}(\alpha, 2\alpha) = 1 + L(\alpha, 2\alpha)$, and lends 1 again to the counterparty until 3α, repeating this until $M\alpha$. At $M\alpha$ she receives $1 + L((M-1)\alpha, M\alpha)$, and pays 1 to the counterparty of the borrowing in 1). The payoff of this strategy is the same as the one of a floating leg: only the regular interest rate payments are left.

Pricing a swap by replication

The cost of the strategy is the value of the rolled lending $Roll_{rf}^{rf}(0; 0, M\alpha; \alpha) = 1$, minus the value of the money borrowed $P^{rf}(0, M\alpha)$, namely

$$1 - P^{rf}(0, M\alpha),$$

and in a market with no arbitrage this must be the fair value of the floating leg,

$$Float_{rf}^{rf}(0; 0, M\alpha; \alpha) = 1 - P^{rf}(0, M\alpha)$$

This is the *fundamental result of the risk-free market*: the floating leg of a swap with last payment at $M\alpha$ can be priced based only on one spot Libor rate as

$$Float_{rf}^{rf}(0; 0, M\alpha) = 1 - \frac{1}{1 + L(0, M\alpha) M\alpha} = \frac{L(0, M\alpha) M\alpha}{1 + L(0, M\alpha) M\alpha}. \qquad (29)$$

Replicating FRA in a riskless Market

The most basic swap is the Forward Rate Agreement (FRA). We consider a FRA with fixing at α and maturity 2α, and we get

$$\begin{aligned} Float_{rf}^{rf}(0;\alpha,2\alpha) &= Float_{rf}^{rf}(0;0,2\alpha) - Float_{rf}^{rf}(0;0,\alpha) \\ &= 1 - P^{rf}(0,2\alpha) - 1 + P^{rf}(0,\alpha) \\ &= P^{rf}(0,\alpha) - P^{rf}(0,2\alpha) \end{aligned}$$

The value of the fixed leg is $P^{rf}(0,2\alpha)\,K\alpha$, leading to

$$FRA_{rf}^{rf}(0;\alpha,2\alpha;K) = P^{rf}(0,\alpha) - P^{rf}(0,2\alpha) - P^{rf}(0,2\alpha)\,K\alpha. \tag{30}$$

Replicating FRA in a riskless Market

$$FRA_{rf}^{rf}(0;\alpha,2\alpha;K) = P^{rf}(0,\alpha) - P^{rf}(0,2\alpha) - P^{rf}(0,2\alpha)\,K\alpha.$$

This implies

$$\begin{aligned} F_{rf}^{rf}(0;\alpha,2\alpha) \quad : & = \frac{1}{\alpha}\left(\frac{P^{rf}(0,\alpha)}{P^{rf}(0,2\alpha)} - 1\right) \\ & = \quad \frac{1}{\alpha}\left(\frac{P_L(0,\alpha)}{P_L(0,2\alpha)} - 1\right) = F_L(0;\alpha,2\alpha) \end{aligned} \tag{31}$$

We see that this model of the market explains why one should expect the first equation of (26) to hold,

$$F\left(t;T,T'\right) \approx F_L\left(t;T,T'\right)$$

FRA as an expectation

We could get the same result without replication, but using the fact that

$$FRA_{rf}^{rf}\left(0;\alpha,2\alpha;K\right)=\mathbb{E}\left[D\left(0,2\alpha\right)\left(\left(L\left(\alpha,2\alpha\right)-K\right)\alpha\right)\right]$$

We can use change of numeraire to eliminate the stochastic discount factor. We redefine the forward measure $\mathbb{Q}^{rf2\alpha}$ as the one associated to the risk-free bond $P^{rf}\left(t,2\alpha\right)$ (distinction is not irrelevant in a multicurve framework). We get

$$FRA_{rf}^{rf}\left(0;\alpha,2\alpha;K\right)=P^{rf}\left(0,2\alpha\right)\mathbb{E}^{rf2\alpha}\left[\left(\left(L\left(\alpha,2\alpha\right)-K\right)\alpha\right)\right].$$

The equilibrium value of K can now also be written as

$$F_{rf}^{rf}\left(0;\alpha,2\alpha\right)=\mathbb{E}^{rf2\alpha}\left[L\left(\alpha,2\alpha\right)\right].$$

Change of Measure

$$F_{rf}^{rf}\left(0;\alpha,2\alpha\right)=\mathbb{E}^{rf2\alpha}\left[L\left(\alpha,2\alpha\right)\right].$$

But under this model

$$L\left(\alpha,2\alpha\right)=L^{rf}\left(\alpha,2\alpha\right)=\frac{1}{\alpha}\left(\frac{P^{rf}\left(\alpha,\alpha\right)-P^{rf}\left(\alpha,2\alpha\right)}{P^{rf}\left(\alpha,2\alpha\right)}\right)$$

so it is the value at α of

$$F_{L}\left(t;\alpha,2\alpha\right)=\frac{1}{\alpha}\left(\frac{P^{rf}\left(t,\alpha\right)-P^{rf}\left(t,2\alpha\right)}{P^{rf}\left(t,2\alpha\right)}\right)$$

which can be seen as *the price of a tradable asset divided by the numeraire* of $\mathbb{Q}^{rf2\alpha}$. This is a martingale under $\mathbb{Q}^{rf2\alpha}$ and we get

$$F_{rf}^{rf}\left(0;\alpha,2\alpha\right)=\mathbb{E}^{rf2\alpha}\left[L\left(\alpha,2\alpha\right)\right]=\frac{1}{\alpha}\left(\frac{P^{rf}\left(0,\alpha\right)}{P^{rf}\left(0,2\alpha\right)}-1\right). \tag{32}$$

A useful remark

Notice that what we have really proved is that

$$\mathbb{E}^{rf2\alpha}\left[L^{rf}\left(\alpha,2\alpha\right)\right]=\frac{1}{\alpha}\left(\frac{P^{rf}\left(0,\alpha\right)}{P^{rf}\left(0,2\alpha\right)}-1\right) \tag{33}$$

but, with the assumption (28) typical of a risk-free market, this corresponds to a result on the expectation of Libor.

Basis Swaps in a Riskless Market

Now we consider a basis swap where two parties exchange two floating legs, both with last payment at T, based on two different payment tenors, α and α', $\alpha < \alpha'$, plus a spread Z added to the α leg. In this setting both floating legs have value

$$Float_{rf}^{rf}(0;0,T;\alpha) = Float_{rf}^{rf}\left(0;0,T;\alpha'\right) = 1 - P^{rf}(0,T)$$

since the fundamental result (29) shows clearly that the value of a floating leg does not depend on its frequency. The spread Z setting the price to zero is

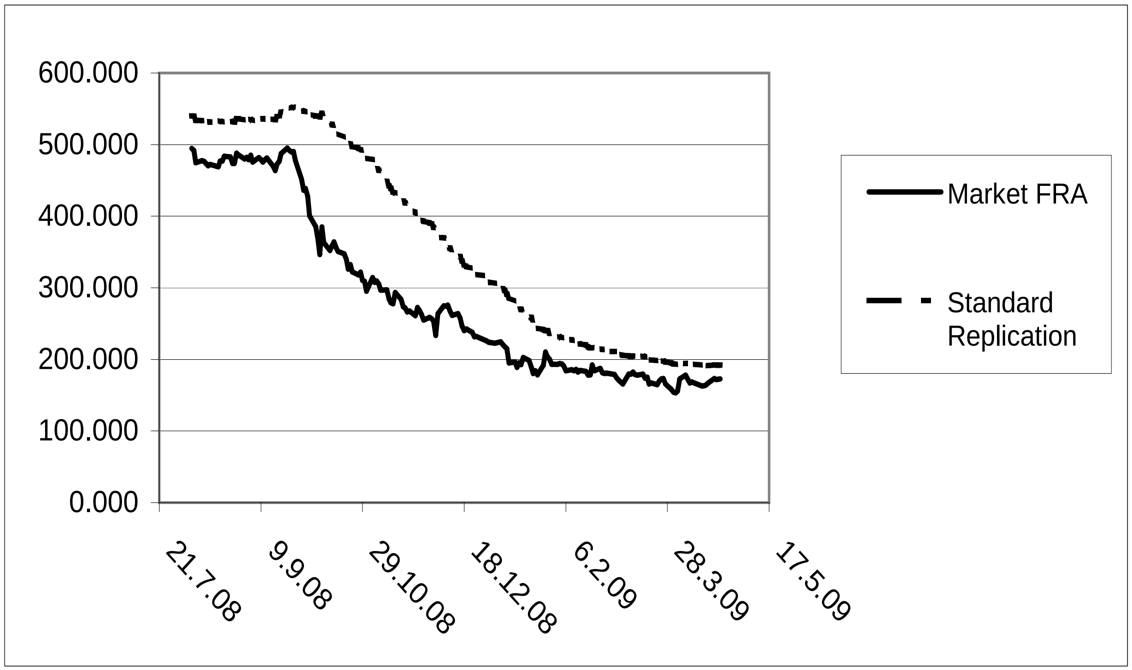
$$B_{rf}^{rf}\left(0;\alpha/\alpha';T\right) = 0.$$

One could indeed replicate a basis swap borrowing and lending the same amount at Libor, a fair deal by definition. Thus we see also why one would expect the second part of (26) to hold.

Thus, based on the above replication arguments, standard financial theory expects this forward rate $F_L\left(t;T,T'\right)$ to be equal to $F\left(t;T,T'\right)$, and B to be zero. This was almost true in the past but it is not anymore:

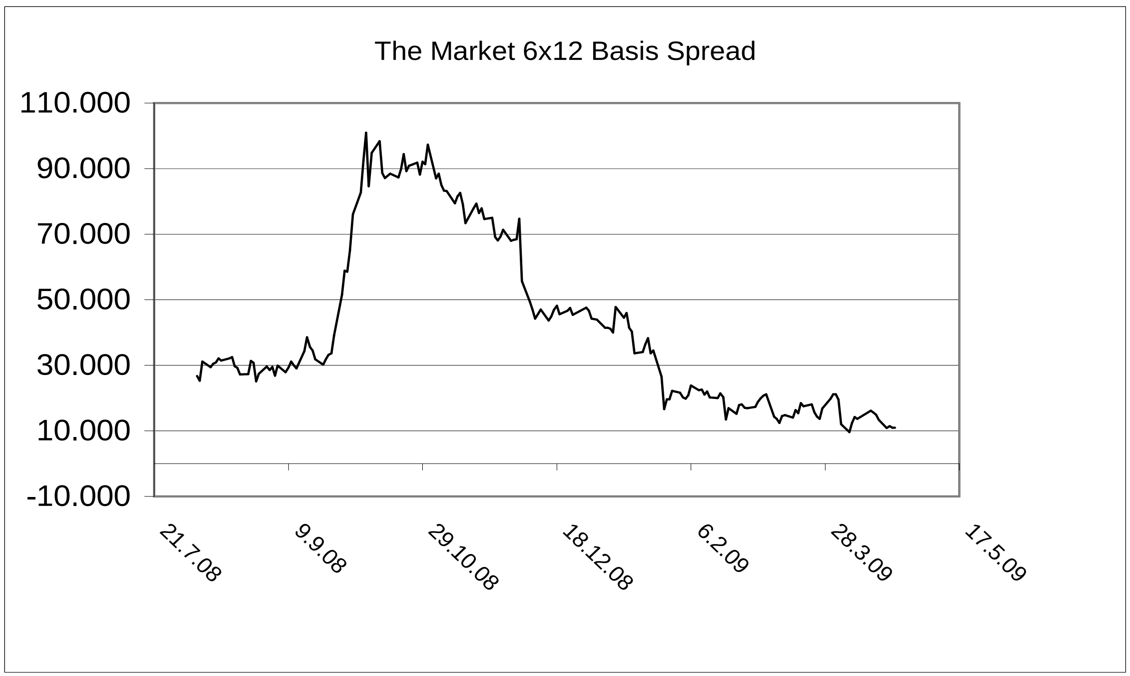
A large, positive gap

Below we can see in more detail both F_L and F in the second half of 2008 and first half of 2009.



The explosion of Basis Spreads

Analogously, the Basis swap spreads widened from very few basis points to much larger values after the crisis. From August 2008 to April 2009, the Basis swap spread to exchange 6 Month Libor with 12 Month Libor over 1 year was strongly positive and averaged 40bps, as we see in Figure 3.



The Rates Market when banks can default

When banks can default, each bank has its own risk of default. The defaultable bond with maturity T issued by a counterparty A has a price

$$P^{dA}(0,T) = \mathbb{E}\left[1_{\{\tau^A > T\}} D(0,T)\right]$$

where τ^A is the default time of A. The rate $L^{dA}(0,T)$ that applies to a loan to counterparty A, for a unit notional, needs to satisfy

$$1 = \mathbb{E}\left[1_{\{\tau^A > T\}} D(0,T)\left(1 + L^{dA}(0,T)\right)\right]$$

leading to

$$\begin{aligned} 1 &= P^{dA}(0,T)\left(1 + L^{dA}(0,T)\right), \\ P^{dA}(t,T) &= \frac{1}{1 + L^{dA}(t,T)} \end{aligned}$$

Notice we have assumed zero recovery for simplicity

The Rates Market when banks can default

In this context, the meaning of the Libor rate $L(t,T)$ is not at easy as in a risk-free market, where there was a unique rate applying to all counterparties. Let $\mathbb{L}_t$ be the set of Libor banks. Each bank $A \in \mathbb{L}_t$ will borrow at a different rate $L^{dA}(t,T)$. The Libor rate is the trimmed average of the $L^{dA}(t,T)$ across all Libor banks, a quantity more complex to model or to replicate than a single rate.

Before the credit crunch, risk of default was not null, and yet the rules of a risk-free market were safely used. Can we understand how this was possible? Are there some assumptions - maybe unrealistic now but common before the crisis - that allow us to simplify the representation of the interest rate market even if there is risk of default? The two assumptions we make are *homogeneity* and *stability*, explained in the following.

Homogeneity Assumption

Homogeneity, or **All Libors were created equal** We say that the set $\mathbb{L}_t$ of Libor banks is *homogeneous* when we can define X_t as the generic Libor bank at t and we have

$$L^{dA}(t,T) = L^{dX_t}(t,T) \qquad \forall A \in \mathbb{L}_t$$

This is the mathematical extremization of the idea, very common before the crisis, that all banks that participate to the interbank market have a similar risk of default. "Libor contains a few basis point of bank vs bank counterparty risk" was a popular description that follows from this idea. Notice that an obvious consequence of X_t being a Libor bank at t is that $\tau^{X_t} > t$. When the credit risk is homogeneous we do not need anymore to differentiate between counterparties and we can write

$$P^{dX_t}(t,T) = \frac{1}{1 + L^{dX_t}(t,T)}.$$

and the Libor quotes have a simple meaning

$$\begin{aligned} L(t,T) &= L^{dX_t}(t,T) \\ P_L(t,T) &= P^{dX_t}(t,T) \end{aligned} \tag{34}$$

Stability Assumption

Stability Assumption, or **Libor today will be Libor for ever** We say that the set $\mathbb{L}_t$ of Libor banks is *stable* when

$$A \in \mathbb{L}_t, \tau^A > T \Longrightarrow A \in \mathbb{L}_T.$$

Homogeneity with stability corresponds to assuming that the future Libor rate at any date $T > t$ will always coincides with the future borrowing rate of the Libor counterparty chosen at t, if the latter has not defaulted. Namely

$$L(t,T) \overset{by\ Hom}{=} L^{dX_t}(t,T) \overset{by\ Stab}{=} L^{dX_0}(t,T), \quad \tau^{X_0} > t$$

This identification between the indexing of interest rate derivatives and the rate of the counterparty is the same feature we have in a risk-free market. Thanks to this, if we consider the same swap as above, it is still possible to evaluate it by replication, at least from the perspective of an investor which is another bank in the Libor market. We start by the replication of a defaultable swap.

Replicating a defaultable swap

The investor can

1) borrow $P^{dX_0}(0, M\alpha)$ until $M\alpha$. Since the investor is a Libor bank, his credit risk is expressed by $L^{dX_0}(0, M\alpha)$ and it is fair for her to borrow $P^{dX_0}(0, M\alpha)$ and give back 1 at $M\alpha$ (in case of no default).

2) lend 1 to another Libor counterparty until α. At α the investor receives $1 + L^{dX_0}(0, \alpha) = 1 + L(0, \alpha)$, and lends 1 again to the same counterparty until 2α. At 2α the investor receives $1 + L^{dX_0}(\alpha, 2\alpha) = 1 + L(\alpha, 2\alpha)$, and lends 1 again to the same counterparty until 3α, repeating this until $M\alpha$ (in case of no default).

This is similar to the replication strategy in the risk-free market. The cost of the strategy is given by the value of the rolled lending $Roll_{dX_0}^{dX_0}(0; 0, M\alpha; \alpha) = 1$ minus the value $P^{dX_0}(0, M\alpha)$ of the money borrowed, namely

$$1 - P^{dX_0}(0, M\alpha)$$

Does this strategy has the same value of a swap floating leg?

Replicating a defaultable swap

This strategy involves receiving a stream of floating payments equal to those of the floating leg, and subject to the same risk of default of a floating leg. Additionally it involves also two capital payments, that do not necessarily cancel out. Thus its value is the value of the floating leg plus the value of the two capital payments:

$$
\begin{aligned}
1 - P^{dX_0}(0, M\alpha) &= Float_{dX_0}^{dX_0}(0; 0, M\alpha; \alpha) + \\
+\mathbb{E}\left[D(0, M\alpha)\, 1_{\{\tau^{Count} > M\alpha\}}\right] &- \mathbb{E}\left[D(0, M\alpha)\, 1_{\{\tau^{Inv} > M\alpha\}}\right].
\end{aligned}
$$

If both the investor and the α-rolled lending counterparty are Libor banks at 0, the strategy has value

$$
\begin{aligned}
1 - P^{dX_0}(0, M\alpha) &= Float_{dX_0}^{dX_0}(0; 0, M\alpha; \alpha) + P^{dX_0}(0, M\alpha) - P^{dX_0}(0, M\alpha) \\
&= Float_{dX_0}^{dX_0}(0; 0, M\alpha; \alpha).
\end{aligned}
$$

Replicating a defaultable swap

$$Float_{dX_0}^{dX_0}(0;0,M\alpha;\alpha) = 1 - P^{dX_0}(0,M\alpha)$$

We have again the fundamental result of risk-free market: the floating leg of a swap with last payment at $M\alpha$ can be priced based only on a Libor rate as

$$Float_{dX_0}^{dX_0}(0;0,T;\alpha) = 1 - \frac{1}{1 + L(0,M\alpha)\,M\alpha} = \frac{L(0,M\alpha)\,M\alpha}{1 + L(0,M\alpha)\,M\alpha}.$$

Replicating a defaultable swap

We consider now a defaultable swap floating leg lasting one single period, for example fixing first at α and ending at 2α. It has a discounted payoff

$$D\left(0,2\alpha\right)1_{\left\{\tau^{X_0}>2\alpha\right\}}L\left(\alpha,2\alpha\right)\alpha. \tag{35}$$

We price it first by replication, following the above derivation: it is the difference between two spot-starting legs as above. We find

$$\begin{aligned} Float_{dX_0}^{dX_0}\left(0;\alpha,2\alpha;\alpha\right) &= P^{dX_0}\left(0,\alpha\right)-P^{dX_0}\left(0,2\alpha\right)= \\ &= P_L\left(0,\alpha\right)-P_L\left(0,2\alpha\right) \end{aligned} \tag{36}$$

The fixed leg of this swap is composed of one single K payment at 2α, which is defaultable too, therefore it has value

$$P^{dX_0}\left(0,2\alpha\right)K\alpha=P_L\left(0,2\alpha\right)K\alpha.$$

Collateralized FRA

In the risk-free market this one-period swap was considered equivalent to a forward rate agreement. If this was the case also in the defaultable market we could conclude again

$$\begin{aligned} F_{dX_0}^{dX_0}(0;\alpha,2\alpha) &= \frac{1}{\alpha}\left(\frac{P_L(0,\alpha)}{P_L(0,2\alpha)}-1\right)=\frac{1}{\alpha}\left(\frac{1+L(0,\alpha)\,\alpha}{1+L(0,2\alpha)\,2\alpha}-1\right), \\ &= F_L\left(t;T,T'\right) \end{aligned}$$

However, in a defaultable market this product cannot correspond to the real market FRA.

1. For *uncollateralized* derivatives, when a default happens the two legs of a swap are not treated separately, as we did above, but there is a *netting*
2. The second unrealistic feature cancels the first one: the FRA are regularly *collateralized*.

Collateralized FRA

Therefore we cannot price it with the above replication for defaultable swaps, nor with the replication for a *risk-free swap in a risk-free market*. If we follow the risk-free replication the fair rate that the risk-free counterparty pays is $L^{rf}(\alpha, 2\alpha)$, that in a defaultable market does not coincide with the Libor paid by the FRA. We need to price a *risk-free swap in a defaultable market*. We have a risk-free (collateralized) derivative with a defaultable (uncollateralized) underlying. This mismatch leads to the impossibility of a simple exact replication.

We can however price the FRA as the expectation of its payoff, getting

$$FRA^{rf}_{dX_0}(0; \alpha, 2\alpha; K) = \mathbb{E}\left[D(0, 2\alpha)\left((L(\alpha, 2\alpha) - K)\,\alpha\right)\right]$$

that looks analogous to FRA^{rf}_{rf} of the risk-free market, but it is not the same since now we have a different model representation of the market quantity $L(\alpha, 2\alpha)$.

Collateralized FRA

Since $FRA^{rf}_{dX_0}$ is a collateralized product, the discount factor is the same as we had in a risk-free market, and we can use again change of numeraire to move to expectation under the same forward measure $\mathbb{Q}^{rf2\alpha}$ associated to $P^{rf}(t, 2\alpha)$ used for the risk free market,

$$\begin{aligned} FRA^{rf}_{dX_0}(0; \alpha, 2\alpha; K) &= P^{rf}(0, 2\alpha)\,\mathbb{E}^{rf2\alpha}\left[\left(\left(L(\alpha, 2\alpha) - K\right)\alpha\right)\right], \\ F^{rf}_{dX_0}(0; \alpha, 2\alpha) &= \mathbb{E}^{rf2\alpha}\left[L(\alpha, 2\alpha)\right]. \end{aligned}$$

The market approach used so far (Mercurio (2008)) stops here, modelling directly this expectations and without trying to understand what $L(\alpha, 2\alpha)$ is. Here, instead, we investigate this issue, and we try to understand quantitatively the gap existing now between FRA rates and the old replication.

Collateralized FRA

$$\mathbb{E}^{rf2\alpha}\left[L\left(\alpha,2\alpha\right)\right]$$

This is easily solved in a risk-free market because there

$$L\left(\alpha,2\alpha\right)=L^{rf}\left(\alpha,2\alpha\right)=\frac{1}{\alpha}\left(\frac{1}{P^{rf}\left(\alpha,2\alpha\right)}-1\right)=\frac{1}{\alpha}\left(\frac{P^{rf}\left(\alpha,\alpha\right)}{P^{rf}\left(\alpha,2\alpha\right)}-1\right)$$

and $\frac{1}{\alpha}\left(\frac{P^{rf}(t,\alpha)}{P^{rf}(t,2\alpha)}-1\right)$ is a martingale since it is a tradable asset divided by the numeraire $P^{rf}\left(t,2\alpha\right)$. Instead in the risky market we have

$$L\left(\alpha,2\alpha\right)=L^{dX_\alpha}\left(\alpha,2\alpha\right)=\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}\left(\alpha,\alpha\right)}{P^{dX_\alpha}\left(\alpha,2\alpha\right)}-1\right)$$

and $\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(t,\alpha)}{P^{dX_\alpha}(t,2\alpha)}-1\right)$ is not a martingale since it is not a tradable asset divided by the numeraire $P^{rf}\left(\alpha,2\alpha\right)$.

Collateralized FRA

Thus

$$\begin{aligned} FRA^{rf}_{dX_0}(0;\alpha,2\alpha;K) &= \mathbb{E}^{rf2\alpha}\left[\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(\alpha,\alpha)}{P^{dX_\alpha}(\alpha,2\alpha)}-1\right)\right] \\ &\neq \frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(0,2\alpha)}{P^{dX_\alpha}(0,2\alpha)}-1\right). \end{aligned}$$

How can compute the expectation even without the martingale property? Detect a measure Q^N, associated to some numeraire N_t, such that

$$\mathbb{E}^N\left[\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(\alpha,\alpha)}{P^{dX_\alpha}(\alpha,2\alpha)}-1\right)\right] = \frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(0,2\alpha)}{P^{dX_\alpha}(0,2\alpha)}-1\right) = F_L(0;\alpha,2\alpha)$$

and then use the mathematical relations between the expectation of $\frac{P^{dX_\alpha}(\alpha,\alpha)}{P^{dX_\alpha}(\alpha,2\alpha)}$ under Q^N and the one under $\mathbb{Q}^{rf2\alpha}$ provided by the change of numeraire techniques.

Change of numeraire for Collateralized FRA

We will get a convexity adjustment, so that

$$F_{dX_0}^{rf}(0;\alpha,2\alpha)=\mathbb{E}^{rf2\alpha}\left[\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(\alpha,\alpha)}{P^{dX_\alpha}(\alpha,2\alpha)}-1\right)\right]=F_L(0;\alpha,2\alpha)+\mathbf{CA}^{N/rf}.$$

In Appendix 2 we find such a measure Q^N and we compute that the equilibrium rate of a collateralized FRA in a risky world is

$$F_{dX_0}^{rf}(0;\alpha,2\alpha) \;=\; F_L(0;\alpha,2\alpha)\exp\left(-\alpha\rho\sigma^Q\sigma^F\right),$$

where σ^F and σ^Q are respectively the volatilities of $F_{dX_0}^{rf}$ and of the ratio between risky and riskless bonds, and ρ is their correlation. See Appendix 2 for details.

Change of measure for Collateralization

Based on a simple historical estimation on Libor and OIS data, considering the crisis period from July 2007 to May 2009, **and a maturity of six months** $(\alpha = 6m)$

$$\sigma^F = 12\%, \sigma^Q = 0.7\%, \rho = 6\%$$

while on a longer period, from January 2003 to May 2009,

$$\sigma^F = 19\%, \sigma^Q = 0.4\%, \rho = 8\%$$

The average value of $F^{rf}_{dX_0}(0; \alpha, 2\alpha)$ was 4.3% in the crisis period and 3.2% in the longer one. This would lead to

$$F^{rf}_{dX_0}(0; \alpha, 2\alpha) = F_L(0; \alpha, 2\alpha) + \begin{cases} \text{-0.02}bp & \text{(based on crisis period)} \\ \text{-0.01}bp & \text{(based on longer period)} \end{cases}$$

Change of measure for Collateralization

Such a convexity adjustment is a number lower in absolute value than one basis point. This is in line with market experience on convexity adjustments for short maturities. Thus we expect

$$F^{rf}_{dX_0}(0;\alpha,2\alpha) \approx F_L(0;\alpha,2\alpha) = \frac{1}{\alpha}\left(\frac{P_L(0,\alpha)}{P_L(0,2\alpha)} - 1\right).$$

The estimation and the assumptions are very rough, but the resulting adjustment is so low - for this $\alpha = 6m$ maturity - that it appears difficult that some changes in the estimated parameters could explain a difference that - for this $\alpha = 6m$ maturity - has been almost as large as 200bps. With homogeneity and stability we cannot explain market anomalies, not even through the measure mismatch.

A useful remark

Notice one thing that will be useful later. What we have actually computed is that

$$\mathbb{E}^{rf2\alpha}\left[L^{dX_0}(\alpha, 2\alpha)\right] \approx \frac{1}{\alpha}\left(\frac{P^{dX_0}(0,\alpha)}{P^{dX_0}(0,2\alpha)} - 1\right) \tag{37}$$

This is a result on the *expectation of the future spread of the current counterparty*. It has been considered *a result on the expectation of Libor* $L(t,T)$ only thanks to homogeneity and stability.

Bridging the gap between Basis and FRA

How can we apply the above results on collateralized FRAs to Basis swaps? Do we need to repeat the above treatment for Basis swaps? Not really. In fact also Basis swaps are collateralized derivatives. Thanks to this we can write a precise relationship between FRA rates and Basis spreads which is model independent. The value of a Basis swap is

$$\begin{aligned} & BS_{?}^{rf}\left(0;\alpha/2\alpha;2\alpha;Z\right) = \mathbb{E}_0\left[D\left(0,\alpha\right)\alpha L(0,\alpha) + D\left(0,2\alpha\right)\alpha L(\alpha,2\alpha) + \right. \\ & \left. -D\left(0,2\alpha\right)2\alpha\left(L(0,2\alpha)-Z\right)\right]. \\ = \quad & P^{rf}\left(0,\alpha\right)\alpha L(0,\alpha) + \mathbb{E}_0\left[D\left(0,2\alpha\right)\alpha L(\alpha,2\alpha)\right] + \\ & -P^{rf}\left(0,2\alpha\right)2\alpha\left(L(0,2\alpha)-Z\right) \end{aligned}$$

Only the component

$$\mathbb{E}_0\left[D\left(0,2\alpha\right)\alpha L(\alpha,2\alpha)\right]$$

really involves the expectation of an unknown, model dependent quantity, the rest are deterministic payments.

Bridging the gap between Basis and FRA

Define

$$\tilde{K}(Z) = \left(\frac{1}{P_L(0,2\alpha)} - 1 - 2Z\alpha - \frac{P^{rf}(0,\alpha)}{P^{rf}(0,2\alpha)}\left(\frac{1}{P_L(0,\alpha)} - 1\right)\right)/\alpha, \quad (38)$$

$$BS^{rf}_{?}(0;\alpha/2\alpha;2\alpha;Z) = \mathbb{E}_0\left[D(0,2\alpha)\,\alpha\,(L(\alpha,2\alpha)\right] - P^{rf}(0,2\alpha)\,\tilde{K}(Z)\,\alpha.$$

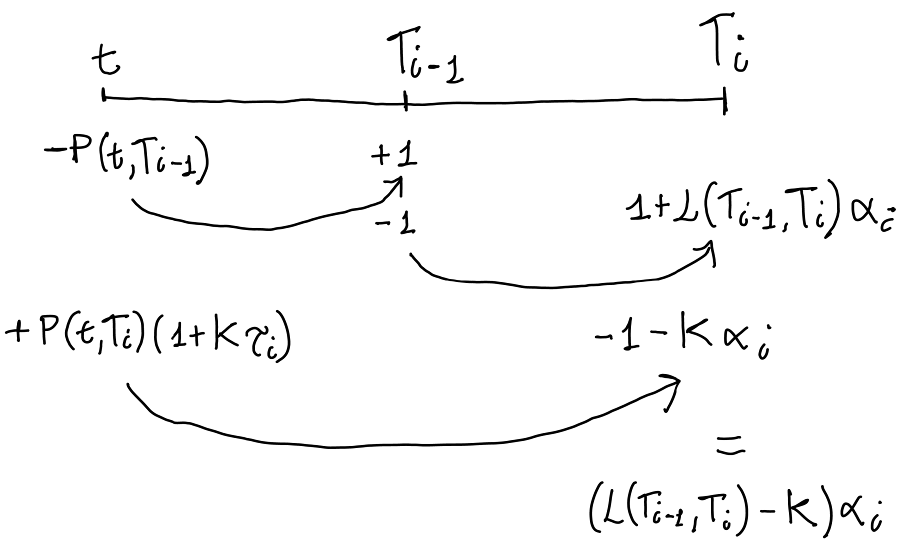
Compare with FRA:

$$\begin{aligned} FRA^{rf}_{?}(0;\alpha,2\alpha;K) &= \mathbb{E}_0\left[D(0,2\alpha)\,\alpha\,(L(\alpha,2\alpha) - K)\right] \\ &= \mathbb{E}_0\left[D(0,2\alpha)\,\alpha L(\alpha,2\alpha)\right] - P^{rf}(0,2\alpha)\,K\alpha. \end{aligned}$$

If one sets $K = \tilde{K}(Z)$ the FRA price is equal to the price of a Basis swap where the spread is set to Z. The first component, the only one that involves a model-dependent expectation of a future Libor rate, is the same for FRA and Basis, so that the two derivatives depend on the same market information.

Bridging the gap between Basis and FRA

Consider the FRA replication strategy. The fixed leg is replicated by a bond with maturity 2α, while the floating leg is replicated by lending α, followed by another lending from α to 2α. One leg has α tenor, the other leg has 2α tenor. FRA is affected by the Basis swap spreads.

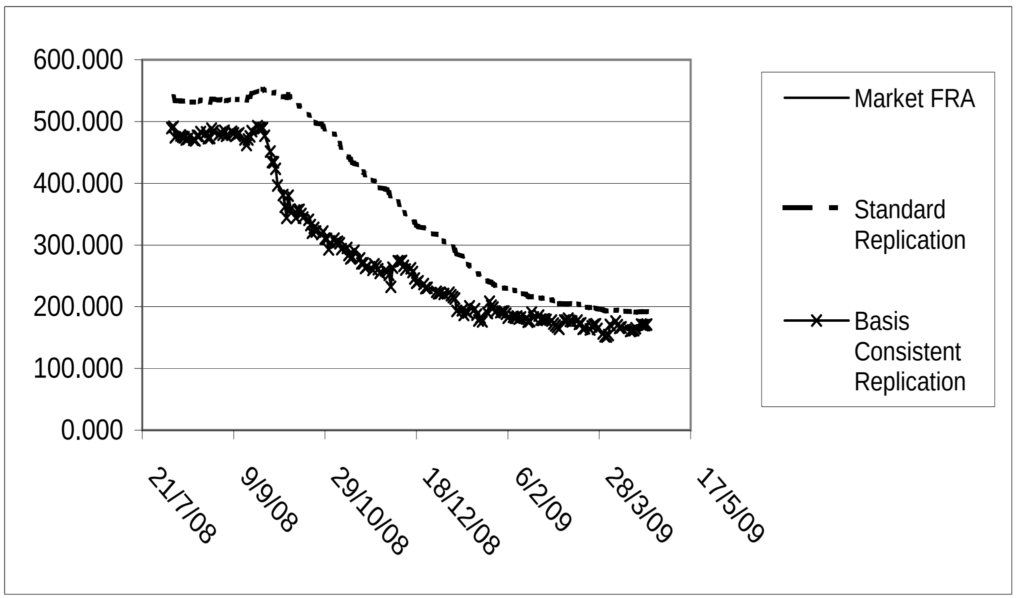


Bridging the gap between Basis and FRA

We have not yet been able to understand how to compute

$$\mathbb{E}_0\left[D\left(0,2\alpha\right)\alpha L(\alpha,2\alpha)\right],$$

but if Z is the market spread that sets to zero the basis price, than in an arbitrage-free market $K=\tilde{K}\left(Z\right)$ should set the FRA price to zero and replicate the equilibrium FRA quote. See below for $\alpha=6m$.



Bridging the gap between Basis and FRA

Viceversa, if one inverts $\tilde{K}(\cdot)$ and sets the basis spread to

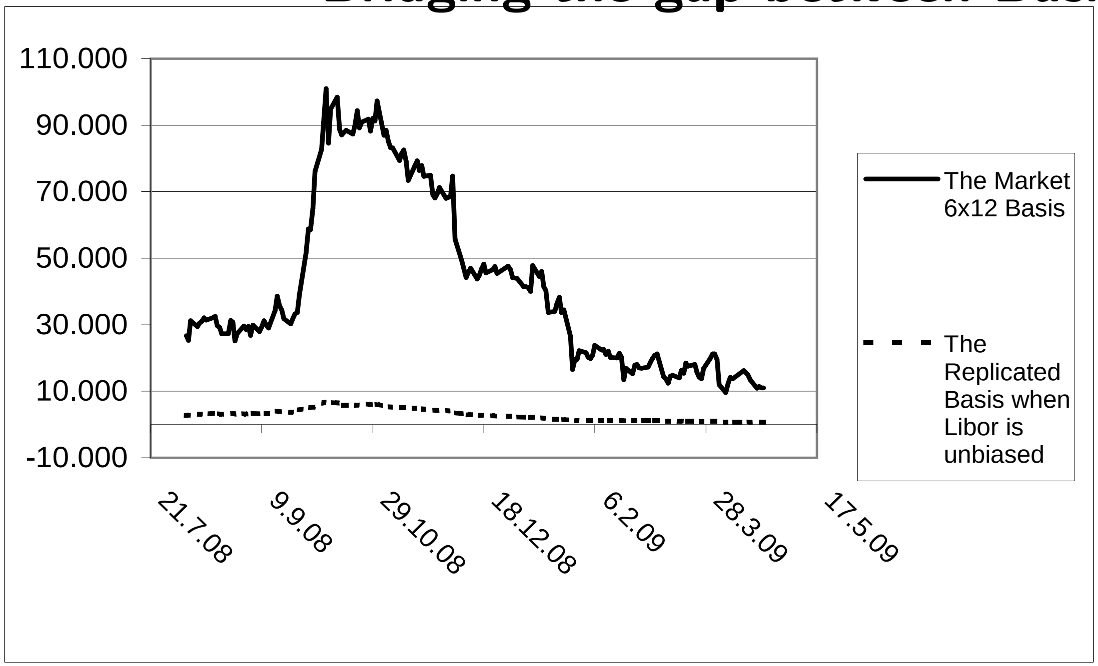
$$\tilde{Z}(K) = \left(\frac{1}{P_L(0,2\alpha)} - \alpha K - \frac{P^{rf}(0,\alpha)}{P^{rf}(0,2\alpha)} \left(\frac{1}{P_L(0,\alpha)} - 1 \right) - 1 \right) / (2\alpha) . \tag{39}$$

the Basis swap will have the same value as the FRA with the fixed rate is set to K. If the equilibrium K is what we have found under homogeneity and stability,

$$F^{rf}_{dX_0}(0;\alpha,2\alpha) \approx \frac{1}{\alpha} \left(\frac{P_L(0,\alpha)}{P_L(0,2\alpha)} - 1 \right) ,$$

this implies a Basis equilibrium spread as

Bridging the gap between Basis and FRA



We are far away from explaining the market pattern. However we have reduced our two problems to only one, and we can concentrate on finding the model explaining either FRA or BS, and then derive the other quote.

Default risk in a stable and homogenous market does not explain FRA anomalies, nor does it explain the recent high Basis. What is the real cause of these market changes?

Common Misconceptions: an option in short lending?

Compare a strategy where I lend for $12m$ with a strategy where I roll over twice a $6m$ lending. In the latter case, after $6m$, if the counterparty default risk has increased, I can stop lending with no cost (at par). If the $12m$ lender wants to stop, instead, will have to unwind at a cost that incorporates the increased risk of default. Thus, some say that *those who roll the $6m$ investment have an option to exit when the credit risk grows, and move to a counterparty with lower risk*, and this may explain the basis. It this true?

Not at all. The $6m$ roller does *not* have an option to exit after $6m$, *because he will always exit at $6m$, including when the default risk of the counterparty has decreased instead of increased*. In case of a decrease of the risk of default of the counterparty, it is the $12m$ lender having an advantage (he receives a yield higher than the actual risk). Thus the expected gain of the $6m$ lender when the counterparty worsens is compensated by its expected loss when the counterparty gets better, and this fact cannot explain the basis. And after $6m$ the roller will not necessarily *move to a counterparty with lower risk*. In fact, such a counterparty will pay a lower yield, and if the investor is looking for yield he may move to an even more risky counterparty. This is normal in the market.

Explaining the Basis. A more realistic market

The above explanations are not totally off the mark, but miss a crucial aspect: the basis does not regard two strategies with different frequencies of payment, but two derivatives which are indexed to Libor rates with different frequencies. The mechanism for the basis must be looked for there, in the Libor panel and market that determine the future level of Libor.

The market with homogeneity and stability used so far is the mathematical representation of a Libor market where all banks have similar risk of default, and this risk of default is not volatile. In this situation we can safely consider negligible the probability that banks exit from the Libor market. This implies that we can take a bank which is today a member of the Libor market and consider its spread a good representation of the Libor rate also for the future.

Let us now suppose that risk of default grows and becomes more volatile, as it happened with the burst of the subprime crisis. Now there are many ways in which this bank can cease being a good representation of Libor in the future..

Banks can exit the Libor panel

Recall the Libor Panel rules:
"*The banks represented on the panels are the most active in the cash markets and have the highest credit ratings*"
"*The BBA is committed to reviewing the Panels at least twice annually*"
When the credit standing of a bank in the Libor panel worsens too much, a bank can even be excluded by the Libor panel and replaced with a better one. This does not happen very often, but it is paradigmatic: clearly, in this case the spread of the originally Libor counterparty will not influence the future Libor rate, that will turn out to be lower than the rate of the counterparty, now cast out of the panel.

Banks can exit the deposit Libor market

A bank can exit the Libor market, even if it does not exit the Libor panel. This was the case for many banks during the subprime crisis. The Libor market is the market where loans are made, unsecured, at a rate which is very similar to the Libor fixings. The Libor fixings are computed by asking banks the rate at which they can borrow unsecured. If a bank has difficulties in borrowing unsecured at quotes close to previous Libor fixings, or even worse it has difficulties to borrow unsecured at all, what is the bank going to answer to the Libor panel? Very likely, it will contribute a fictive rate, that rather than being their true rate for unsecured borrowing is their guess about the rate for unsecured borrowing that applies to those other banks that have no difficulties in getting unsecured borrowing. This is the thesis exposed, exactly for the credit crunch period of interest to us, in Peng et al. (2008), claiming Libor underestimated funding costs.

This is confirmed by the fact that although the crisis saw the *freezing up of the interbank market* (Brunnermeier (2008)) contributions for unsecured lending rates went on reaching the panel with no apparent discontinuity.

Bank can exit the interquartile part of the Libor panel

The bank can remain in the Libor panel and in the Libor market and yet not be anymore a good representative of Libor. In fact Libor is a trimmed average, in the sense that the highest and lowest quartile of the contributions are taken away before averaging; when risk of default is stable and homogenous, this fact is negligible, but when there is more variability this may have a relevance. Our reference bank can turn out to be either better or worse than Libor. Which one the two cases is more relevant? In crisis times, the case with quantitatively higher consequences seems to be when the bank's credit spread is worse than average.

Banks can exit the world of the living banks...

There is one last thing to consider, that can happen for banks even still in the panel. Eventually, if their credit standing worsens too much, unless they are bailed-out by a government, banks can default. This is what happened to Lehman. In the following we do not consider default explicitly, but we implicitly treat it as one the events that can bring a bank out of the Libor market when its spreads worsen too much. At times banks default without an enormous increase in the spread, but there is always some correlations between spread rallies and default event. This was the case of Lehman during the credit crunch. Default is one more selection mechanism that creates a bias towards banks with lower risk.

The swap when credit risk is volatile

Therefore, *the expected future spread of Libor is constantly lower than the expected spread in the future of a bank which is now a Libor counterparty*. This means that we have lost the *stability assumption*. And without stability, also *homogeneity* cannot strictly hold. We can assume homogeneity to hold at time 0 for simplicity, but it cannot hold at all times. What is the effect of this conclusion on our approach to replicate the floating leg of a swap? A disruptive effect. The standard replication strategy of a floating leg does not make sense anymore.

The swap when credit risk is volatile

Suppose that the investor willing to replicate the floating leg does as follows:

Lend today 1 for α years to a counterparty X_0 which is today a libor counterparty. After α years it will receive $1 + L^{dX_0}(0, \alpha) = 1 + L(0, \alpha)$. So far the strategy is the same as above, but now the investor needs changing approach.

At α she should check if the counterparty is still a Libor bank, namely one of the banks with the lowest credit risk. If the answer is no, namely $X_0 \notin \mathbb{L}_\alpha$ because it is too risky, she cannot lend the money again to the same counterparty, otherwise she would receive $1 + L^{dX_0}(\alpha, 2\alpha) > 1 + L(\alpha, 2\alpha)$. Even the first fundamental bit of the replication would be wrong, overestimating the level of Libor.

The investor must instead choose another counterparty X_α with a lower credit risk and thus a lower credit spread, so that $X_\alpha \in \mathbb{L}_\alpha$ and the investor will receive $1 + L^{dX_\alpha}(\alpha, 2\alpha) = 1 + L(\alpha, 2\alpha)$. This process, involving every α an assessment of the risk of counterparty with possible substitution with a counterparty with lower credit spread, must be repeated until $M\alpha$. The credit quality of the counterparty is assessed, and improved if necessary, every α years.

The FRA

Now suppose we are in a market FRA, a collateralized product where there is only one stochastic payment involved, the one at α. Compared to the risk-free or stable-risk cases in which one did not change the counterparty, lending always to X_0, now at α we may need to move to a counterparty better than X_0 in case X_0 exits the Libor panel or market. Therefore the α-payment $L^{dX_\alpha}(\alpha, 2\alpha)$ will be lower than $L^{dX_0}(\alpha, 2\alpha)$ that we had in the previous market models, so that also the fixed rate needed to set the two legs in equilibrium will be lower, giving an intuitive understanding, that we will verify analytically below, of why we can now explain

$$F^{rf}_{dX_0 X_\alpha}(0; \alpha, 2\alpha) \ll \frac{1}{\alpha}\left(\frac{P_L(0, \alpha)}{P_L(0, 2\alpha)} - 1\right).$$

The Basis Swap

It becomes important to understand what changes when, from a swap with α tenor ($\frac{1}{\alpha}$ frequency) like the one above, we move to different tenors. If the tenor of the interest rate payment in the floating leg was 2α (thus with payments at $2\alpha, 4\alpha, 6\alpha \ldots$) the process of assessment and improvement of the counterparty credit quality would be done less often. The investor will be exposed to all possible worsening of the counterparty credit that can happen along 2α years. Thus the 2α-year Libor leg involves more credit risk than the α-year Libor leg, and this higher credit risk is embed in higher rates paid. If instead the tenor of the interest rate payment in the floating leg was $\frac{\alpha}{2}$ (thus with payments at $\frac{\alpha}{2}, \alpha, \frac{3}{2}\alpha \ldots$), this assessment and improvement of the counterparty should be done every $\frac{\alpha}{2}$ years, reducing the risk of default even more than in the α-year Libor leg.

In the α leg the Libor rates are higher since they embed higher credit spreads, a compensation for the higher counterparty risk. But with collateralization the two legs have the same null counterparty risk. For reaching equilibrium a spread Z needs to be added to the $\frac{\alpha}{2}$ leg,

$$B\left(t; \alpha/\alpha'; T\right) \gg 0$$

The overnight rate

There is another important implication of this representation of the interest rate market. In the replication of a floating leg, the more frequent the payments, the lower the credit risk, until reaching the shortest tenor available in the market: the overnight. To replicate a leg with overnight payments, one should assess and if necessary improve the credit quality of the counterparty everyday, leading to a leg involving an extremely low level of counterparty risk. Who is admitted to the overnight lending market? Those borrowers whose risk of default is negligible over the next day. This explains why the overnight quote is usually considered to be default-free. Since only banks with a negligible default risk over the term of the lending are admitted to this market, its quote represent a very good approximation of a risk-free interest rate.

A Model for the Basis: abandoning Stability

We want to go beyond intuition and model Libor rates when credit risk is high and volatile, so that homogeneity holds only at time 0 and stability does not hold at all, meaning that banks can go out of the Libor market. With homogeneity and stability we had,

$$L(\alpha, 2\alpha) \overset{by\ Hom}{=} L^{dX_\alpha}(\alpha, 2\alpha) \overset{by\ Stab}{=} L^{dX_0}(\alpha, 2\alpha).$$

Now we have

$$L(\alpha, 2\alpha) \leq L^{dX_0}(\alpha, 2\alpha).$$

We will have to verify if we can go on with this identification or not, by checking if $L^{dX_0}(\alpha, 2\alpha)$ satisfy the requirements for X_0 being a Libor bank.

A Model for the Basis: abandoning Stability

Notice that both $L(\alpha, 2\alpha)$ and $L^{dX_0}(\alpha, 2\alpha)$ are given by the risk-free rate $L^{rf}(\alpha, 2\alpha)$ plus a spread,

$$\begin{aligned} L(\alpha, 2\alpha) &=: L^{rf}(\alpha, 2\alpha) + S(\alpha, 2\alpha), \\ L^{dX_0}(\alpha, 2\alpha) &=: L^{rf}(\alpha, 2\alpha) + S^{dX_0}(\alpha, 2\alpha). \end{aligned}$$

What do we know about the rates and spreads involved in our computations? We have computed that

$$\begin{aligned} \mathbb{E}^{rf2\alpha}\left[L^{rf}(\alpha, 2\alpha)\right] &= \frac{1}{\alpha}\left(\frac{P^{rf}(0,\alpha)}{P^{rf}(0,2\alpha)} - 1\right), \\ \mathbb{E}^{rf2\alpha}\left[L^{dX_0}(t, T)\right] &\approx \frac{1}{\alpha}\left(\frac{P^{dX_0}(0,\alpha)}{P^{dX_0}(0,2\alpha)} - 1\right) \\ \mathbb{E}^{rf2\alpha}\left[S^{dX_0}(\alpha, 2\alpha)\right] &\approx \frac{1}{\alpha}\left(\frac{P^{rf}(0,\alpha)}{P^{rf}(0,2\alpha)} - \frac{P^{dX_0}(0,\alpha)}{P^{dX_0}(0,2\alpha)}\right). \end{aligned}$$

A Model for the Basis: abandoning Stability

We introduce the variable

$$S^{dX_0}(t;\alpha,2\alpha)=\frac{1}{\alpha}\left(\frac{P^{rf}(t,\alpha)}{P^{rf}(t,2\alpha)}-\frac{P^{dX_0}(t,\alpha)}{P^{dX_0}(t,2\alpha)}\right),$$

that we call the forward spread of X_0 and that has the property of being a martingale (having neglected the small convexity adjustment) and that

$$S^{dX_0}(\alpha;\alpha,2\alpha)=S^{dX_0}(\alpha,2\alpha)$$

We know that $S(\alpha,2\alpha)\leq S^{dX_0}(\alpha,2\alpha)$ when the counterparty worsens too much.

The Libor mechanism: a simple scheme

A simple scheme to represent the potential refreshment of the counterparty is:

$$\boxed{S(\alpha, 2\alpha) = \begin{cases} S^{dX_0}(\alpha, 2\alpha) & \text{if } S^{dX_0}(\alpha, 2\alpha) \leq S^{Exit} \\ S^{Subst} & \text{if } S^{dX_0}(\alpha, 2\alpha) > S^{Exit} \end{cases}}, \tag{40}$$

where

- S^{Exit} **is the maximum level of the spread for** X_0 **to be still a Libor counterparty at** α
- S^{Subst} **is the spread of Libor in case** $X_0 \notin \mathbb{L}_\alpha$ **because** $S^{dX_0}(\alpha, 2\alpha) > S^{Exit}$

The Libor mechanism: a simple scheme

This is the same as writing

$$
\begin{aligned}
L(\alpha, 2\alpha) &= L^{rf}(\alpha, 2\alpha) \\
&\quad + 1_{\left\{S^{dX_0(\alpha,2\alpha) \le S^{Exit}}\right\}} S^{dX_0}(\alpha, 2\alpha) + 1_{\left\{S^{dX_0(\alpha,2\alpha) > S^{Exit}}\right\}} S^{Subst} \\
&= L^{rf}(\alpha, 2\alpha) \\
&\quad + S^{dX_0}(\alpha, 2\alpha) - 1_{\left\{S^{dX_0(\alpha,2\alpha) > S^{Exit}}\right\}} \left(S^{dX_0}(\alpha, 2\alpha) - S^{Subst}\right) \\
&= L^{dX_0}(\alpha, 2\alpha) - 1_{\left\{S^{dX_0(\alpha,2\alpha) > S^{Exit}}\right\}} \left(S^{dX_0}(\alpha, 2\alpha) - S^{Subst}\right).
\end{aligned}
$$

The Libor mechanism: Exit condition

First we have to make a choice on S^{Exit} , than on S^{Subst}, with

$$S^{Subst} \leq S^{Exit}.$$

We will not aim to make any precise representation of reality; we will propose as an example some extremely simple possible values and we will see if even such simplified assumptions can lead to relevant improvements in explaining the market compared to the old risk-free or stable-risk-of-default approaches.

The level S^{Exit} is the level over which a Libor bank is considered an underperformer. Thus S^{Exit} can be close to the current expectation of $S^{dX_0}(\alpha, 2\alpha)$, so that *a counterparty will be excluded from Libor banks if it performs worse than expected*:

$$\boxed{S^{Exit} = S^{dX_0}(0; \alpha, 2\alpha) := \mathbb{E}^{rf2\alpha}\left[S^{dX_0}(\alpha, 2\alpha)\right]}$$

The Libor mechanism: substitution condition

One simple choice for S^{Subst}: the counterparty X_0 is excluded by Libor because its spread is higher than today expectation by an amount

$$S^{dX_0}(\alpha, 2\alpha) - S^{dX_0}(0; \alpha, 2\alpha) > 0;$$

in this case we assume that the prototypical counterparty that has been included in Libor did as well as X_0 did badly, thus we move to Libor *whose spread is lower than today expectation by the same amount* $S^{dX_0}(\alpha, 2\alpha) - S^{dX_0}(0; \alpha, 2\alpha)$, so that

$$S^{Subst} = S^{dX_0}(0; \alpha, 2\alpha) - \left[S^{dX_0}(\alpha, 2\alpha) - S^{dX_0}(0; \alpha, 2\alpha)\right]$$

$$\boxed{S^{Subst} = 2S_{dX_0}(0; \alpha, 2\alpha) - S^{dX_0}(\alpha, 2\alpha)}.$$

An Option in Libor quotes

This leads to

$$\begin{aligned} L(\alpha, 2\alpha) &= L^{dX_0}(\alpha, 2\alpha) - 1_{\left\{S^{dX_0(\alpha,2\alpha)>S^{Exit}}\right\}}\left(S^{dX_0}(\alpha, 2\alpha) - S^{Subst}\right) \\ &= L^{dX_0}(\alpha, 2\alpha) - 2\left(S^{dX_0}(\alpha, 2\alpha) - S^{dX_0}(0; \alpha, 2\alpha)\right)^+ \\ &= L^{dX_0}(\alpha, 2\alpha) - 2\left(S^{dX_0}(\alpha; \alpha, 2\alpha) - S^{dX_0}(0; \alpha, 2\alpha)\right)^+ . \end{aligned}$$

We are ready to solve our problem:

$$\mathbb{E}^{rf2\alpha}[L(\alpha, 2\alpha)] = F_L(0; \alpha, 2\alpha) - 2\mathbb{E}^{rf2\alpha}\left[\left(S^{dX_0}(\alpha; \alpha, 2\alpha) - S^{dX_0}(0; \alpha, 2\alpha)\right)^+\right]$$

An Option in Libor quotes

$$\mathbb{E}^{rf2\alpha}\left[L\left(\alpha,2\alpha\right)\right]=F_L\left(0;\alpha,2\alpha\right)-2\mathbb{E}^{rf2\alpha}\left[\left(S^{dX_0}\left(\alpha;\alpha,2\alpha\right)-S^{dX_0}\left(0;\alpha,2\alpha\right)\right)^+\right]$$

If we assume that $S^{dX_0}\left(t,\alpha,2\alpha\right)$ evolves as a geometric brownian motion

$$dS_{dX_0}\left(t,\alpha,2\alpha\right)=S^{dX_0}\left(t,\alpha,2\alpha\right)\sigma_\alpha dW_\alpha^P\left(t\right),$$

we have the simple option formula

$$\begin{aligned}\mathbb{E}^{rf2\alpha}\left[L\left(\alpha,2\alpha\right)\right]&=F_L\left(0;\alpha,2\alpha\right)-\\ &2BlackCall\left(S^{dX_0}\left(0;\alpha,2\alpha\right),S^{dX_0}\left(0;\alpha,2\alpha\right),\sigma_\alpha\sqrt{\alpha}\right).\end{aligned}\tag{41}$$

An Option in Libor quotes

Like old $F_L(0;\alpha,2\alpha)$, this formula gives expected future Libor based on current Libor. But now it includes a refreshment mechanism that can explain the market patterns of FRA. This shows precisely why in this model the FRA equilibrium rate is lower than $F_L(0;\alpha,2\alpha)$, since

$$F^{rf}_{dX_0,X_\alpha}(0;\alpha,2\alpha) = \mathbb{E}^{rf2\alpha}\left[L(\alpha,2\alpha)\right] \leq F_L(0;\alpha,2\alpha),$$

with the inequality collapsing to an equality only if $S^{dX_0}(0;\alpha,2\alpha)$ is zero, that would bring us to the risk-free market, or if the credit volatility is zero, that would bring us back to the stable-default-risk market.

An Option in Libor quotes

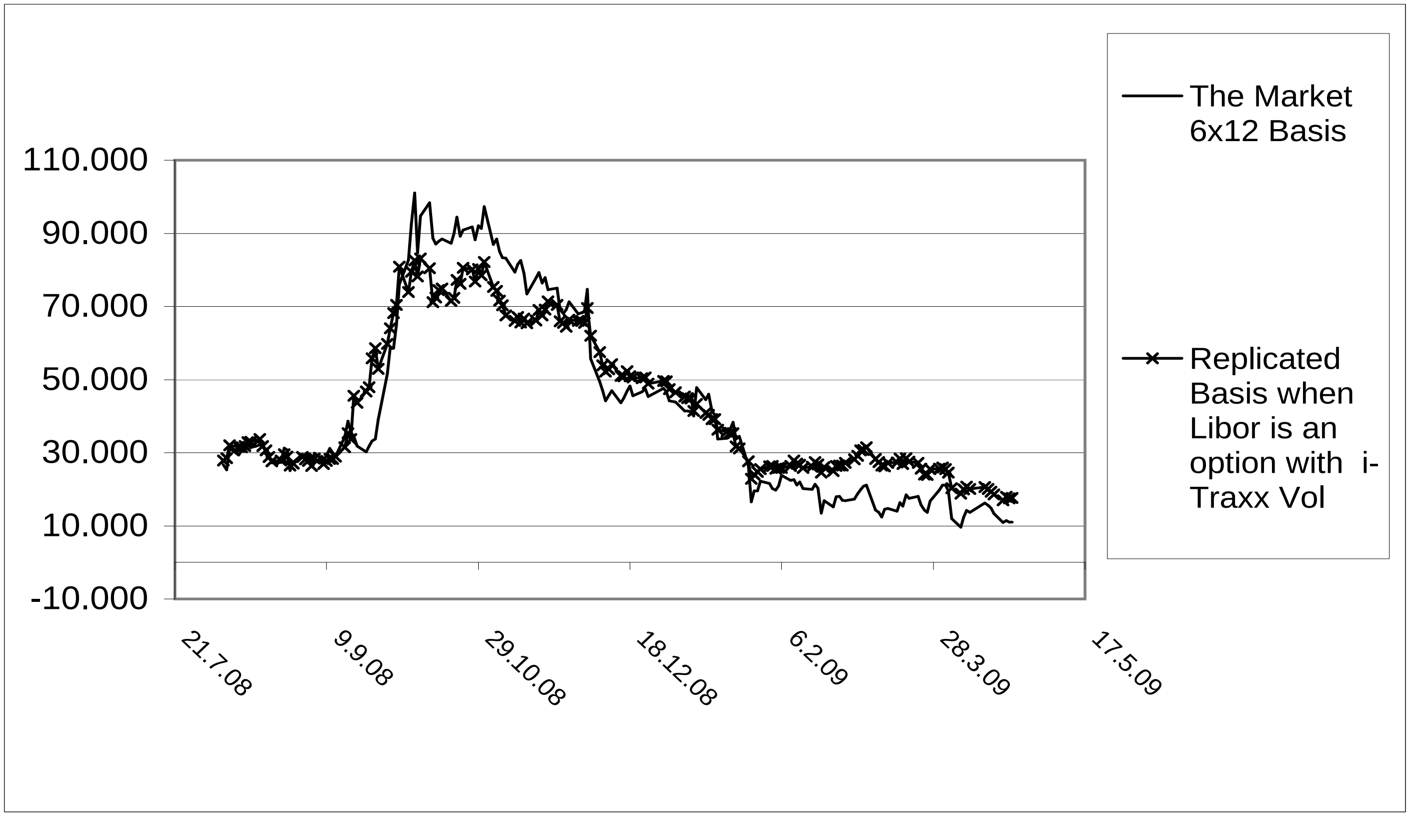
This formulation can also explain the non-null basis spread. But we do not want to explain only the sign of the Basis spread, but also its magnitude, so we will test our formula on real market quotes of Basis swaps. We consider the Basis swap where as usual $\alpha = 6m$. We have that

$$\begin{aligned} B^{rf}_{dX_0,X_\alpha}(0;\alpha/2\alpha;2\alpha) &= \frac{1}{2}\left(\frac{1}{P_L(0,\alpha)} - 1\right) \times \\ &\times \left(F_L(0;\alpha,2\alpha) - F^{rf}_{rf}(0;\alpha,2\alpha)\right) + \\ &- BlackCall\left(S^{dX_0}(0;\alpha,2\alpha), S^{dX_0}(0;\alpha,2\alpha), \sigma_\alpha\sqrt{\alpha}\right) \end{aligned}$$

As a proxy for the credit spread volatility σ_α we use the ATM implied volatility of the i-Traxx Index spread, the average credit spread of the most liquid entities in the Euro market. The risk-free rates come instead from OIS quotes.

Explaining the Basis

The simple formula based uniquely on Euribor and OIS data, with a credit volatility input, yields a good replication of the historical behaviour of the traded $6m/12m$ Basis:



Future Research

Various points would deserve further investigation:

- The volatility input is based on the most liquid European CDS, but we need a volatility referring only to Financials. The volatility chosen can explain underestimation during the pitch of the crisis, and overestimation more recently.
- The spread dynamics can be more realistic (jumps).
- The options could be a collar rather than a cap since Libor is a trimmed average.
- The model is very simple. The choice on the substitute counterparty is very optimistic. A less optimistic choice on this, coupled with allowing liquidity reasons for exiting Libor, may lead to similar results with easier interpretation.

Remarks on the Interest Rate Puzzle: OIS and FX analogy

Appendix 1: OIS for long maturity risk-free rates

Our explanation of the lack of risk in the overnight market seems to imply that, in a market where default risk is relevant for all players when considering a term longer than one day, risk-free quotes can only be available for a maturity of one day. The market overcomes this problem by quoting Overnight Indexed Swaps (OIS). For short maturities they are simple instruments which provide for the exchange at maturity T of two amounts of money. One counterparty pays a fixed amount KT, the other counterparty pays an amount which is computed at T as the compounding of the overnight rates that fixed in the period from 0 to T: If in the period there have been n business days, indicated ad t_i, $i = 0, \ldots, n$, with $t_0 = 0$ and $t_n = T$, the amount paid is

$$\left(\prod_{i=1}^{n} \left[1 + L\left(t_{i-1}, t_i\right)\left(t_i - t_{i-1}\right)\right] - 1\right).$$

Appendix 1: OIS for long maturity risk-free rates

The rate K is fixed so as to guarantee the contract is fair. The contract is collateralized and risk-free, and the rates paid are, according to the above reasoning, rates that embed no risk. Thus the equilibrium K is found by enforcing

$$\begin{aligned} \mathbb{E}\left[D(0,T)\left(\prod_{i=1}^{n}\left[1+L(t_{i-1},t_i)(t_i-t_{i-1})\right]-1\right)\right] &= \mathbb{E}\left[D(0,T)KT\right] \\ \mathbb{E}\left[D(0,T)\prod_{i=1}^{n}\left[1+L(t_{i-1},t_i)(t_i-t_{i-1})\right]\right] &= P^{rf}(0,T)(1+KT) \end{aligned}$$

Appendix 1: OIS for long maturity risk-free rates

We know that in a risk free world one can obtain a risk-free payoff $\prod_{i=1}^{n}\left[1+L\left(t_{i-1},t_i\right)\left(t_i-t_{i-1}\right)\right]$ at T simply by rolling lending an initial amount of 1. Thus we need to have

$$\mathbb{E}\left[D\left(0,T\right)\prod_{i=1}^{n}\left[1+L\left(t_{i-1},t_i\right)\left(t_i-t_{i-1}\right)\right]\right]=1$$

which leads to an equilibrium level for K given by

$$OIS\left(0,T\right)=\left(\frac{1}{P^{rf}\left(0,T\right)}-1\right)\frac{1}{T}$$

This means that $OIS\left(0,T\right)$ corresponds to the rate that we called $L^{rf}\left(0,T\right)$, expressing the time-value of money in the absence of credit risk. Thus it provides a risk-free rate even when the market is fraught with credit risk,

$$L^{rf}\left(0,T\right)=OIS_M\left(0,T\right).$$

Appendix 2: Probability and Measure for a defaultable market

Starting from

$$\mathbb{E}\left[D(0,2\alpha)\,1_{\left\{\tau^{X_0}>2\alpha\right\}}L(\alpha,2\alpha)\right],$$

we look for a convenient numeraire N_t to express the expectation of $L(\alpha,2\alpha)$ in a simple form. The purpose is to express

$$F_{dX_0}^{rf}(0;\alpha,2\alpha)=\mathbb{E}^{rf2\alpha}\left[L(\alpha,2\alpha)\right]=\mathbb{E}^{rf2\alpha}\left[\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(\alpha,\alpha)}{P^{dX_\alpha}(\alpha,2\alpha)}-1\right)\right]$$

as

$$\mathbb{E}^{N}\left[\frac{1}{\alpha}\left(\frac{P^{dX_\alpha}(\alpha,\alpha)}{P^{dX_\alpha}(\alpha,2\alpha)}-1\right)\right]+\mathbf{CA}^{N/rf}$$

We work under homogeneity and stability.

Change of numeraire for Collateralized FRA

Notice by the way that the convenient numeraire N_t cannot be

$$P^{dX_0}(t, 2\alpha) = \mathbb{E}_t \left[D(t, 2\alpha) \, 1_{\left\{ \tau^{X_0} > 2\alpha \right\}} \right],$$

that may appear a natural choice. In fact a numeraire needs to satisfy two requirements

1. It must be a tradable asset, where tradable asset means loosely "the expectation of a discounted payoff".
2. It must be strictly positive, for its mathematical nature of a "denominator".

Appendix 2: Probability and Measure for a defaultable market

Like in standard intensity models for CDS, we assume that information given by $\mathcal{F}_s$ is divided into subfiltrations defined by

$$\begin{aligned}\mathcal{F}_t &= \mathcal{H}_t \vee \mathcal{J}_t \\ \mathcal{J}_t &= \sigma\left(\{\tau > u\}, u \leq t\right),\end{aligned} \tag{42}$$

where $\mathcal{J}_t$ is the natural filtration of the default time τ of a counterparty, while $\mathcal{H}_t$ is the *no-default information*, information up to t on economic quantities which affect default probability, such as the default-free interest rates, but excluding specific information on happening of default. When default is unpredictable, it is natural to assume that $\mathcal{H}$-martingales are also $\mathcal{F}$-martingales, and $\mathbb{Q}\left(\tau > t | \mathcal{H}_t\right) > 0$.

In this context a defaultable payoff $Y = 1_{\{\tau>t\}}Y$ depending on no-default information and on the default time τ of a given player can be priced using only a default indicator and no-default information,

$$\mathbb{E}\left[\mathbf{Y}|\mathcal{F}_t\right] = \frac{\mathbf{1}_{\{\tau>t\}}}{\mathbb{Q}\left(\tau > t|\mathcal{H}_t\right)}\mathbb{E}\left[\mathbf{Y}|\mathcal{H}_t\right]. \tag{43}$$

A general definition of Libor when banks can default

In this context the price of a defaultable bond at $s > t$ is

$$P^{dX_t}(s,T) = \mathbb{E}\left[D(s,T)\,\mathbf{1}_{\left\{\tau^{X_t}>T\right\}}|\mathcal{F}_s\right] = \frac{1_{\left\{\tau^{X_t}>s\right\}}\mathbb{E}\left[D(s,T)\,\mathbf{1}_{\left\{\tau^{X_t}>T\right\}}|\mathcal{H}_s\right]}{\mathbb{Q}\left(\tau^{X_t}>s|\mathcal{H}_s\right)},$$

This expression of the price allows us also to give for the first time an explicit representation of the Libor bond. We know that under homogeneity and stability $P_L(s,T)$ coincides with $P^{dX_t}(s,T)$ *provided that we have chosen an X_t which is still alive at s.* This corresponds to say that we must have $1_{\left\{\tau^{X_t}>s\right\}} = 1$, so that

$$P_L(s,T) = \frac{\mathbb{E}\left[D(s,T)\,\mathbf{1}_{\left\{\tau^{X_t}>T\right\}}|\mathcal{H}_s\right]}{\mathbb{Q}\left(\tau^{X_t}>s|\mathcal{H}_s\right)},$$

where t can be any date such that $t < s$. Clearly we are assuming that there will always be at least one counterparty alive at s to be taken as a reference.

A general definition of Libor when banks can default

This Libor bond never goes to zero and we also have a unit value at maturity

$$P_L(T,T) = \frac{\mathbb{E}\left[\mathbf{1}_{\left\{\tau^{X_t}>T\right\}}|\mathcal{H}_T\right]}{\mathbb{Q}\left(\tau^{X_t}>T|\mathcal{H}_T\right)} = \frac{\mathbb{Q}\left(\tau^{X_t}>T|\mathcal{H}_T\right)}{\mathbb{Q}\left(\tau^{X_t}>T|\mathcal{H}_T\right)} = 1.$$

Is this Libor bond $P_L(s,T)$ a valid numeraire? Now we have property 2), since it is strictly positive. But we have lost property 1), since this is not a tradable asset, in fact it cannot be interpreted as the expectation of a future discounted payoff. If we take only its numerator, we have the expectation of a future discounted payoff. Jamshidian (2004) shows that in fact the quantity

$$\tilde{P}^{dX_t}(s,2\alpha) = \mathbb{E}\left[D(s,2\alpha)\,\mathbf{1}_{\left\{\tau^{X_t}>2\alpha\right\}}|\mathcal{H}_s\right]$$

is a valid numeraire when dealing with $\mathcal{H}_s$ expectations.

The equilibrium Forward rate

After some computations we get

$$\begin{aligned} Float_{dX_0}^{dX_0}\left(t;\alpha,2\alpha;\alpha\right)/\alpha &= \frac{\alpha\mathbb{E}\left[D\left(t,2\alpha\right)\mathbb{Q}\left(\tau^{X_t}>2\alpha|\mathcal{H}_{2\alpha}\right)L(\alpha,2\alpha)|\mathcal{H}_t\right]}{\mathbb{Q}\left(\tau^{X_t}>t|\mathcal{H}_t\right)} \\ &= P_L\left(t,2\alpha\right)\alpha\mathbb{E}^{dX_0 2\alpha}\left[L(\alpha,2\alpha)|\mathcal{H}_t\right]. \end{aligned}$$

Now are interested in computing

$$\mathbb{E}^{dX_0 2\alpha}\left[\frac{1}{\alpha}\left(\frac{1}{P_L\left(\alpha,2\alpha\right)}-1\right)|\mathcal{H}_t\right].$$

Martingales and default probabilities

$$
\begin{aligned}
& \mathbb{E}^{dX_0 2\alpha}\left[\frac{1}{\alpha}\left(\frac{1}{P_L(\alpha,2\alpha)}-1\right)|\mathcal{H}_t\right] \\
= \quad & \mathbb{E}^{dX_0 2\alpha}\left[\frac{1}{\alpha}\left(\frac{\mathbb{Q}\left(\tau^{X_t}>\alpha|\mathcal{H}_\alpha\right)}{\tilde{P}^{dX_t}(\alpha,2\alpha)}-1\right)|\mathcal{H}_t\right] \\
= \quad & \mathbb{E}^{dX_0 2\alpha}\left[\frac{1}{\alpha}\left(\frac{\tilde{P}^{dX_t}(\alpha,\alpha)}{\tilde{P}^{dX_t}(\alpha,2\alpha)}-1\right)|\mathcal{H}_t\right]
\end{aligned}
$$

and it can be proved that $\frac{1}{\alpha}\left(\frac{\tilde{P}^{dX}t(s,\alpha)}{\tilde{P}^{dX}t(s,2\alpha)}-1\right)$ is a martingale under the $\tilde{P}^{dX_t}(s,2\alpha)$ pricing measure and equals $\frac{1}{\alpha}\left(\frac{P_L(s,\alpha)}{P_L(s,2\alpha)}-1\right)$, leading to

$$
F_{dX_0}^{dX_0}(t;\alpha,2\alpha)=\mathbb{E}^{dX_0 2\alpha}\left[L(\alpha,2\alpha)|\mathcal{H}_t\right]=\frac{1}{\alpha}\left(\frac{1+L(t,\alpha)\alpha}{1+L(t,2\alpha)2\alpha}-1\right).
$$

We already knew this by replication, but the mathematical derivation has got a couple of advantages.

The same Payoff, Many risks

The less practical advantage is to show how to deal consistently with two different bonds, the Libor bond $P_L(t,T)$ and the risk-free bond $P^{rf}(t,T)$, which embed different risk in spite of the fact that they give the same non-defaultable payoff of 1 at T. This apparent *arbitrage* is explained by the fact that only the risk-free bond is a tradable asset, and therefore can be chosen as a numeraire, while $P_L(t,T)$ is not a tradable asset. This explain how it can embed default risk without defaulting, but cannot be used as a numeraire. Thus, differently from what one would expect in standard change of numeraire, when above we have to compute

$$\mathbb{E}^{dX_0 2\alpha}\left[L(\alpha,2\alpha)|\mathcal{H}_t\right] = \mathbb{E}^{dX_0 2\alpha}\left[\frac{1}{\alpha}\left(\frac{1}{P_L(\alpha,2\alpha)} - 1\right)|\mathcal{H}_t\right]$$

we do not treat $P_L(\alpha,2\alpha)$ at the denominator as the numeraire and we do not treat 1 at numerator as the value at maturity of a risk-free bond $P^{rf}(\alpha,\alpha) = 1$.

The same Payoff, Many risks

In spite of this, we are able to prove the martingale property since

$$P_L(\alpha, 2\alpha) = \frac{\tilde{P}^{dX_0}(\alpha, 2\alpha)}{\mathbb{Q}(\tau^{X_0} > \alpha | \mathcal{H}_\alpha)}$$

therefore we have to compute the expectation of a new ratio

$$\mathbb{E}^{dX_0 2\alpha}[L(\alpha, 2\alpha)|\mathcal{H}_t] = \mathbb{E}^{dX_0 2\alpha}\left[\frac{1}{\alpha}\left(\frac{\tilde{P}^{dX_0}(\alpha, \alpha)}{\tilde{P}^{dX_0}(\alpha, 2\alpha)} - 1\right)|\mathcal{H}_t\right]$$

where the denominator is actually the numeraire and the numerator is the value at maturity of the $\mathcal{H}_t$ expectation of a discounted payoff (a ‘tradable asset’).

More Realism: Collateralization

From a practical point of view, the interest of the above is that it gives a framework to price a collateralized FRA. The equilibrium rate of a collateralized FRA is

$$F^{rf}_{dX_0}(0;\alpha,2\alpha) = \mathbb{E}^{rf2\alpha}[L(\alpha,2\alpha)]$$

We have seen that

$$\mathbb{E}^{dX_0 2\alpha}[L(\alpha,2\alpha)|\mathcal{H}_t] = \frac{1}{\alpha}\left(\frac{P_L(t,\alpha)}{P_L(t,2\alpha)} - 1\right) = F^{dX_0}_{dX_0}(t;\alpha,2\alpha) \tag{44}$$

where the difference with FRA is in the expectation, taken under the measure associated to $\tilde{P}^{dX_0}(t,2\alpha)$ rather than $P^{rf}(t,2\alpha)$. We exploit the fact that the rate $F^{dX_0}_{dX_0}(t;\alpha,2\alpha)$ is a martingale under the $\tilde{P}^{dX_0}(t,2\alpha)$ measure and has the property

$$F^{dX_0}_{dX_0}(\alpha;\alpha,2\alpha) = L^{dX_0}(\alpha,2\alpha) = L(\alpha,2\alpha)\,.$$

More Realism: Collateralization

The equilibrium rate $F^{rf}_{dX_0}(0;\alpha,2\alpha)$ differs from (44) by the result of a change of measure (convexity adjustment),

$$\begin{aligned} F^{rf}_{dX_0}(0;\alpha,2\alpha) &= \mathbb{E}^{rf2\alpha}\left[L(\alpha,2\alpha)\right] = \mathbb{E}^{dX_0 2\alpha}\left[L(\alpha,2\alpha)\right] + CA(0,\alpha) \\ &= F^{dX_0}_{dX_0}(t;\alpha,2\alpha) + CA(0,\alpha) \end{aligned}$$

The convexity adjustment depends on the volatility of the numeraire ratio

$$\frac{\tilde{P}^{dX_0}(t,\alpha)}{P^{rf}(t,\alpha)}.$$

Cross-Currency Analogy explained

For computing the dynamics of $F_{dX_0}^{dX_0}(t;\alpha,2\alpha)$ under the measure associated to the risk-free bond $P^{rf}(t,2\alpha)$ we need the dynamics of

$$\frac{\tilde{P}^{dX^0}(t,2\alpha)}{P^{rf}(t,2\alpha)} = \frac{\mathbb{Q}(\tau > t|\mathcal{H}_t)\,P_L(t,2\alpha)}{P^{rf}(t,2\alpha)} =: Q(t,2\alpha)$$

Cross-currency analogy (Bianchetti (2008)) explained by credit: $Q(t,2\alpha)$ has the same form as the forward exchange rate $X(t,2\alpha)$ that regulates cross-currency change of numeraire,

$$X(t,2\alpha) = \frac{x(t)\,P_F(t,2\alpha)}{P_D(t,2\alpha)}$$

where $P_F(t,2\alpha)$, $P_D(t,T)$ are respectively the foreign and domestic bond prices, and $x(t)$ is the spot exchange rate that converts the foreign bond in domestic currency.

Change of measure for Collateralization

We write for simplicity

$$F_{dX_0}^{dX_0}(t;\alpha,2\alpha) =: F(t;\alpha,2\alpha)$$

and we assume lognormal dynamics

$$\begin{aligned} dF(t;\alpha,2\alpha) &= \sigma^F(t)F(t;\alpha,2\alpha)\ dW_F^{dX_0}(t) \\ dQ(t,2\alpha) &= \sigma^Q(t)Q(t,2\alpha)\ dW_Q^{rf}(t) \end{aligned}$$

Change of measure for Collateralization

If we introduce the correlation ρ between $dW_{F_{2\alpha}}(t)$ and $dW_{Q_{2\alpha}}(t)$,

$$\begin{aligned} dW^{dX_0}(t) &= dW^{rf} - \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \begin{bmatrix} 0 \\ \sigma^Q(t) \end{bmatrix} dt, \\ dW_{F_{2\alpha}}^{dX_0 2\alpha}(t) &= dW_{F_{2\alpha}}^{rf2\alpha} - \rho\sigma_{2\alpha}^Q(t)dt, \end{aligned}$$

$$\begin{aligned} dF\,(t;\alpha,2\alpha) &= \sigma^F(t)F\,(t;\alpha,2\alpha)\;\; dW_F^{dX_0}(t) \\ &= -\rho\sigma^Q(t)\sigma^F(t)F\,(t;\alpha,2\alpha)\,dt + \sigma^F(t)F\,(t;\alpha,2\alpha)\,dW_F^{rf} \end{aligned}$$

Under flat parameters

$$\begin{aligned} F_{dX_0}^{rf}\,(0;\alpha,2\alpha) &= \mathbb{E}^{rf2\alpha}\left[L\,(\alpha,2\alpha)\right] = \mathbb{E}^{rf2\alpha}\left[F_{dX_0}^{dX_0}\,(\alpha;\alpha,2\alpha)\right], \\ &= F_{dX_0}^{dX_0}\,(0;\alpha,2\alpha)\exp\left(-\alpha\rho\sigma^Q\sigma^F\right) \\ &= F_L\,(0;\alpha,2\alpha)\exp\left(-\alpha\rho\sigma^Q\sigma^F\right), \end{aligned}$$

Credit and Liquidity. How funding enters into pricing

Outline

- Computing the funding component of a price
- Liquidity charge without double counting with credit charge
- DVA as a funding benefit
- The role of the Basis on the cost of funding
- Managing liquidity risk: value vs carry

Credit and Funding

Funding liquidity and credit risk are closely related. Banks compute a spread for funding costs that includes a compensation for their own risk of default. However, interactions between the two are still poorly understood. We look for a unified consistent framework for liquidity and credit risk.

We show first that a naive application of the standard approach to including funding costs by modifying the discounting rate, when it is put in place together with the standard approach for the computation of CVA (credit value adjustment) and DVA (debt value adjustment) leads to double-counting of assets that can be realized only once. This can be avoided by taking explicitly into account, in the valuation of a derivative, also the funding strategy and the effect of a default on this strategy.

Our results have an effect also on the cost of funding to be charged to borrowers, as we show later on.

The simplest derivative

One entity, that we call B (*borrower*), commits to pay a fixed amount K at time T to a party L (*lender*). This is the derivative equivalent of a zero-coupon bond issued by B or a loan from L to B.

This way we will be able to compare the results with well-established market practice. Here the net lender and the net borrower never exchange role, and the liquidity cost/benefit arises clearly from the premium. Stochastic option payoffs are similar, while in bilateral payoffs like swaps the situation is different and liquidity needs arise stochastically during the life of the deal.

CDS spread

Party X has a recovery rate R_X and that the risk free interest rate that applies to maturity T has a deterministic value r (it is an approximation for the OIS rate). A party X makes funding in the *bond market*, and is a *CDS reference entity*. We have the following information:

1. the instantaneous deterministic CDS spread π_X. Following the standard CDS market intensity model,
$$\pi_X = \lambda_X \, \mathsf{LGD}_X \tag{45}$$
where λ_X is the deterministic default intensity and $\mathsf{LGD}_X = 1 - \mathsf{R}_X$ is the loss given default of entity X. If recovery is null, we have $\mathsf{LGD}_X = 1$ and the CDS spread coincides with λ_X, so that
$$\Pr(\tau_X > T) = e^{-\pi_X T}$$
π_X must reflect the best estimate of the risk-adjusted default probability and recovery of X. We initially assume $\mathsf{R}_X = 0$.

The liquidity basis

2 the cost of funding s_X. For most issuers this is measured in the secondary bond market. It must be intended as a spread over OIS. We take s_X to be instantaneous and deterministic too. We can compute by difference a liquidity basis γ_X such that

$$s_X = \pi_X + \gamma_X$$

Apart from sovereigns, usually $\gamma_X \geq 0$. It is associated with the cost of the liquidity provision that the buyer makes to the issuer when buying the bond. This is dependent on the greater or lower ease with which the bonds of X can be sold in the secondary market, thus γ_X is related to both funding liquidity costs and market liquidity risk.

Standard DVA: is something missing?

Let's start from the market standard for CVA and DVA. The above deal has a premium paid by the lender L that makes the deal fair at time 0

$$P = \mathbb{E}\left[e^{-rT} 1_{\{\tau_B > T\}} K\right] = e^{-rT} \mathbb{Q}[\tau_B > T] K = e^{-rT} e^{-\pi_B T} K$$

The difference between this premium and the risk-free value is,

$$\begin{aligned} \mathsf{CVA}_L &= \mathbb{E}[e^{-rT} 1_{\{\tau_B \le T\}}] K = e^{-rT} \mathbb{Q}[\tau_B \le T] K \\ &= e^{-rT}\left[1 - e^{-\pi_B T}\right] K \end{aligned}$$

so that

$$P = e^{-rT} K - \mathsf{CVA}_L$$

At the same time party B sees the premium as

$$P = e^{-rT} K - \mathsf{DVA}_B$$

with $\mathsf{CVA}_L = \mathsf{DVA}_B$.

Standard DVA: is something missing?

This guarantees the symmetry for the parties to agree on the premium of the deal,

$$P = e^{-rT} e^{-\pi_B T} K, \tag{46}$$

but it does not consider explicitly the value of liquidity. In fact at time 0 party B receives a cash flow from party L equal to P, so while party L has to finance the amount until the maturity of the deal at its funding spread s_L, party B can reduce its funding by P. So party B should see a funding benefit, and party L a financing costs.

Standard DVA plus liquidity: Is something duplicated?

How come that these funding components do not appear in the above valuation? The absence of the funding term for L can indeed be justified by assuming $s_L = 0$. This implies $\pi_L = 0$. However the same assumption cannot be made for B without changing completely the nature of the deal. In fact assuming $s_B = 0$ would imply $\pi_B = 0$, which would cancel the DVA and CVA term. Thus when B is a party with non-negligible risk of default he must have a funding cost given at least by $s_B = \pi_B > 0$. The effect of this funding costs seems to be missing in the above formula. In the next sections we analyze if it is really missing.

To introduce liquidity costs, we follow Piterbarg (2010), that corrects the discounting rate by the funding spread. According to Piterbarg (2010), the funding spread includes the CDS spread (default intensity) of an institution; but he does not consider explicitly default of the two parties.

Standard DVA plus liquidity: Is something duplicated?

We introduce liquidity costs by adjusting the discounting term, but we also introduce defaultability of the payoff, getting for the lender

$$\begin{aligned} V_L &= \mathbb{E}\left[e^{-\left(r+s_L\right)T} K 1_{\{\tau_B > T\}}\right] - P \qquad (47)\\ &= \mathbb{E}\left[e^{-rT} e^{-\gamma_L T} e^{-\pi_L T} K 1_{\{\tau_B > T\}}\right] - P \\ &= e^{-\left(r+\gamma_L+\pi_L+\pi_B\right)T} K - P, \end{aligned}$$

and analogously for the borrower

$$\begin{aligned} V_B &= -\mathbb{E}\left[e^{-\left(r+s_B\right)T} K 1_{\{\tau_B > T\}}\right] + P \qquad (48)\\ &= -\mathbb{E}\left[e^{-rT} e^{-\pi_B T} e^{-\gamma_B T} K 1_{\{\tau_B > T\}}\right] + P \\ &= -e^{-r\,T} e^{-\pi_B T} e^{-\gamma_B T} K e^{-\pi_B T} + P \\ &= -e^{-\left(r+\gamma_B+2\pi_B\right)T} K + P \end{aligned}$$

Standard DVA plus liquidity: Is something duplicated?

To compare this result, including CVA, DVA and liquidity from discounting, with results on DVA obtained previously, it is convenient to reduce ourselves to the simplest situation where L is default free and with no liquidity spread, while B is defaultable and has the minimum funding cost allowed in this case: $s_L = 0$, $s_B = \pi_B > 0$. We have

$$\begin{aligned} P_L &= e^{-rT} e^{-\pi_B T} K \\ P_B &= e^{-r\,T} e^{-2\pi_B\,T} K = e^{-r\,T} e^{-\pi_B\,T} \boxed{\; e^{-\pi_B\,T} \;} K \end{aligned}$$

There are two bizarre aspects in this representation.

1. even with no liquidity spread, two counterparties do not agree on the simplest transaction with default risk. A *day-one profit* should be accounted by all borrowers.
2. there is a duplication of the funding benefit for the party that assumes the liability. If this were correct then a consistent accounting of liabilities at fair value would require pricing zero-coupon bonds by multiplying *twice* their risk-free present value by their survival probabilities.

Solving the puzzle: risky funding

In order to solve the puzzle, we model explicitly the funding strategy. Here companies capitalize and discount money with the risk-free rate r, and then add or subtract credit and funding costs. The above deal has two legs. For the lender L, one is the **deal leg**, with net present value

$$\mathbb{E}\left[-P + e^{-rT}\,\Pi\right]$$

where Π is the payoff at T, including a potential default indicator; the other leg is the **funding leg** with net present value

$$\mathbb{E}\left[+P - e^{-rT}\,F\right]$$

where F is the funding payback at T, including a potential default indicator. When there is no default risk or liquidity cost, this funding leg can be overlooked because

$$\mathbb{E}\left[+P - e^{-rT}\,e^{rT}P\right] = 0.$$

Instead, in the general case the total net present value is

$$V_L = \mathbb{E}\left[-P + e^{-rT}\,\Pi + P - e^{-rT}\,F\right] = \mathbb{E}\left[e^{-rT}\,\Pi - e^{-rT}\,F\right].$$

Risky Funding with DVA for the borrower

The borrower B has a liquidity advantage from receiving the premium P, as it allows him to reduce its funding requirement by an equivalent amount P. The amount P of funding would have generated a negative cashflow at T, when funding must be paid back, equal to

$$-P\, e^{rT}\, e^{s_B T}\, 1_{\{\tau_B > T\}} \tag{49}$$

The outflow equals P capitalized at the cost of funding, times a default indicator $1_{\{\tau_B > T\}}$.

Why do we need to include a default indicator $1_{\{\tau_B > T\}}$? Because in case of default, under the assumption of zero recovery, the borrower does not pay back the borrowed funding and there is no outflow. Thus reducing the funding by P corresponds to receiving at T a positive amount equal to (49) in absolute value,

$$P\, e^{rT}\, e^{s_B T}\, 1_{\{\tau_B > T\}} = P\, e^{rT}\, e^{\pi_B T}\, e^{\gamma_B T}\, 1_{\{\tau_B > T\}}$$

to be added to what B has to pay in the deal. Thus the total payoff at T is

$$1_{\{\tau_B > T\}} P\, e^{rT}\, e^{\pi_B T}\, e^{\gamma_B T} - 1_{\{\tau_B > T\}} K \tag{50}$$

Risky Funding with DVA for the borrower

$$1_{\{\tau_B>T\}} P\, e^{rT}\, e^{\pi_B T}\, e^{\gamma_B T} - 1_{\{\tau_B>T\}} K$$

Taking discounted expectation,

$$\begin{aligned} V_B &= e^{-\pi_B T}\, P\, e^{\pi_B T}\, e^{\gamma_B T} - K\, e^{-\pi_B T}\, e^{-rT} \\ &= P\, e^{\gamma_B T} - K\, e^{-\pi_B T}\, e^{-rT} \end{aligned} \tag{51}$$

Now we have no unrealistic double accounting of default probability. Notice that

$$V_B = 0 \quad \Rightarrow \quad P_B = K\, e^{-\left(r+\gamma_B+\pi_B\right)T}$$

Assume, as above, that $\gamma_B = 0$ so that in this case

$$P_B = K\, e^{-\pi_B T}\, e^{-rT}. \tag{52}$$

and compare with (46). Taking into account the probability of default in the valuation of the funding benefit removes any liquidity advantage for the borrower.

Risky funding with CVA for the lender

If the lender pays P at time 0, he incurs a liquidity cost. In fact he needs to finance (borrow) P until T. At T, L will give back the borrowed money with interest, but only if he has not defaulted. Otherwise he gives back nothing, so the outflow is

$$P\,e^{rT}\,e^{s_L T}\,1_{\{\tau_L > T\}} = P\,e^{rT}\,e^{\gamma_L T}\,e^{\pi_L T}\,1_{\{\tau_L > T\}}.$$

The total payoff at T is therefore

$$-P\,e^{rT}\,e^{\gamma_L T}\,e^{\pi_L T}\,1_{\{\tau_L > T\}} + K\,1_{\{\tau_B > T\}}. \tag{53}$$

Taking discounted expectation

$$V_L = -P\,e^{\gamma_L T} + K\,e^{-rT}\,e^{-\pi_B T}$$

Risky funding with CVA for the lender

$$V_L = -P\, e^{\gamma_L T} + K\, e^{-rT}\, e^{-\pi_B T}$$

The condition that makes the deal convenient for the lender is

$$V_L = 0 \quad \Rightarrow \quad P_L = K\, e^{-\left(r+\gamma_L+\pi_B\right)T} \tag{54}$$

The lender, when he computes the value of the deal taking into account all future cashflows as they are seen from the counterparties, does not include a charge to the borrower for that component π_L of the cost of funding which is associated with his own risk of default. This is canceled by the fact that funding is not given back in case of default.

The conditions for market agreement

For reaching an agreement in the market we need

$$V_L \geq 0,\ V_B \geq 0$$

which implies

$$\begin{aligned} P_L &\geq P \geq P_B \\ K\, e^{-rT}\, e^{-\gamma_L T}\, e^{-\pi_B T} &\geq P \geq K e^{-rT} e^{-\gamma_B T}\, e^{-\pi_B T} \end{aligned} \tag{55}$$

Thus an agreement can be found whenever

$$\boxed{\gamma_B \geq \gamma_L}$$

This solves the puzzle, and shows that, if we only want to guarantee a positive *expected* return from the deal, the liquidity cost that needs to be charged to the counterparty of an uncollateralized derivative transaction is just the liquidity basis, rather than the bond spread or the CDS spread. But we will also look at the possible *realized* cashflows.

Positive recovery extension

If we relax the assumption of zero recovery, the discounted payoff for the borrower is now

$$
\begin{aligned}
& 1_{\{\tau_B > T\}} e^{-rT} \, P \, e^{\pi_B T} \, e^{\gamma_B T} \, e^{rT} \\
& + 1_{\{\tau_B \le T\}} e^{-r\,\tau_B} \, \mathsf{R}_B \, e^{-r\,(T-\tau_B)} \, P \, e^{\pi_B T} \, e^{\gamma_B T} e^{rT} \\
& - 1_{\{\tau_B > T\}} e^{-rT} \, K \\
& - 1_{\{\tau_B \le T\}} e^{-r\,\tau_B} \, \mathsf{R}_B \, e^{-r\,(T-\tau_B)} \, K
\end{aligned}
$$

Simplifying the terms and taking the expectation at 0 we obtain

$$
\begin{aligned}
V_B &= \mathbb{Q}\{\tau_B > T\} \, P \, e^{\pi_B T} \, e^{\gamma_B T} + \mathbb{Q}\{\tau_B \le T\} \, e^{-r\,T} \, \mathsf{R}_B \, P \, e^{\pi_B T} \, e^{\gamma_B T} e^{rT} \\
&\quad - \mathbb{Q}\{\tau_B > T\} \, e^{-rT} \, K - \mathbb{Q}\{\tau_B \le T\} \, \mathsf{R}_B \, e^{-r\,T} \, K \\
&= [1 - \mathsf{L}\textsf{GD}_B \, \mathbb{Q}\{\tau_B \le T\}] \left(P \, e^{\pi_B T} \, e^{\gamma_B T} - e^{-r\,T} \, K \right). \qquad (56)
\end{aligned}
$$

Positive recovery extension

$$V_B = [1 - \mathsf{L}\text{GD}_B \, (1 - \mathbb{Q}\{\tau_B > T\})] \left(P\, e^{\pi_B T}\, e^{\gamma_B T} - e^{-r\,T}\, K\right)$$

We can write the first order approximation

$$1 - e^{-\pi_B\,T} \approx \mathsf{L}\text{GD}_B \left(1 - e^{-\lambda_B\,T}\right)$$

which allows us to approximate (56) as

$$\begin{aligned} V_B &\approx e^{-\pi_B T} \left(P\, e^{\pi_B T}\, e^{\gamma_B T} - e^{-r\,T}\, K\right) \\ &= P\, e^{\gamma_B T} - e^{-\pi_B T}\, e^{-r\,T}\, K \end{aligned}$$

The previous result is recovered as a first order approximation in the general case of positive recovery rate. Similar arguments apply to the value of the claim for L.

The accounting view for the borrower

What is the meaning of DVA? Are we really taking into account a benefit that will be concretely observed just in case of our default? In this section we show what happens if the borrower pretends to be default-free. In this case the premium P paid by the lender gives B a reduction of the funding payback at T corresponding to a cashflow at T

$$Pe^{rT}e^{s_B T},$$

where there is no default indicator because B is treating itself as default-free. This cashflow must be compared with the payout of the deal at T, which is $-K$ again without indicator. Thus the total payoff at T is

$$e^{rT}e^{s_B T} - K \tag{57}$$

By discounting to zero we obtain an accounting value

$$V_B = P\,e^{s_B T} - K\,e^{-rT}$$

The accounting view for the borrower

$$V_B = P\, e^{s_B T} - K\, e^{-rT}$$

This yields an accounting breakeven premium P_B for the borrower *equal to the previous breakeven*

$$P_B = K\, e^{-rT} e^{-\gamma_B T} e^{-\pi_B T}, \tag{58}$$

Also in this case the borrower B recognizes on its liability a funding benefit that takes into account its own market risk of default π_B, plus additional liquidity basis γ_B. It does not matter how one divides the funding spread into credit spread and liquidity basis.

The accounting view for the lender

We showed that borrower's valuation does not change if he considers himself default free. Do we have a similar property also for the lender? Not at all. He gets a different breakeven premium, because

$$P_L = K\, e^{-rT}\, e^{-\gamma_L T}\, e^{-\pi_B T}$$

depends crucially on γ_L.

The accounting view for the lender

This is not the only difference between the situation of the borrower and the lender's. The borrower's net payout at maturity T is non-negative in all states of the world if we keep $P \geq P_B$, although the latter condition was designed only to guarantee that the *expected* payout is non-negative. For the lender instead the payout at maturity is

$$-P\, e^{rT}\, e^{\gamma_L T}\, e^{\pi_L T}\, 1_{\{\tau_L > T\}} + K\, 1_{\{\tau_B > T\}}.$$

The condition for the non-negativity of the expected payout of the lender does not imply the non-negativity of all realized payouts, even when no one defaults. We additionally require

$$\pi_L \leq \pi_B. \tag{59}$$

If the latter condition is not satisfied, the lender is exposed to negative carry even if the deal is, on average, convenient to him.

Always pretending to be default-free

Liquidity shortages when no one defaults can be excluded by imposing for each deal $\pi_L \leq \pi_B$, or, with a solution working for whatever deal with whatever counterparty, by working as if the lender was default-free. In this case the condition for the convenience of the deal based on expected cashflows becomes

$$P \ \leq \ K\, e^{-rT}\, e^{-s_L T}\, e^{-\pi_B T} = K\, e^{-rT}\, e^{-\gamma_L T}\, e^{-\pi_L T} e^{-\pi_B T}$$

that implies the non-negativity of all payouts. Assuming ourselves to be default-free is consistent with Piterbarg (2010), and in fact it yields for a bank X that pretends to be default-free a breakeven

$$P_B = P_X = e^{-s_X T} e^{-rT} K$$

when it is in a borrower position, while when it is in a lender position the breakeven premium (**with a risk-free counterparty**) will be given by

$$P_L = e^{-s_X T} e^{-rT} K = P_B = P_X.$$

A difficult market...

But on the other hand, for general counterparties with non-null credit risk and liquidity costs, the lender's assumption to be default-free makes a market agreement very difficult, since in this case the agreement

$$K\, e^{-rT}\, e^{-\gamma_B T}\, e^{-\pi_B T} \leq P \leq K\, e^{-rT}\, e^{-\gamma_L T}\, e^{-\pi_L T} e^{-\pi_B T}$$

implies

$$\boxed{\gamma_B \geq \gamma_L + \pi_L}$$

rather than

$$\boxed{\gamma_B \geq \gamma_L}$$

A difficult market...

In a market where everyone treats himself as default-free and counterparties as defaultable, a party wants to fund itself at a spread that includes only its own CDS

$$(\gamma_B + \pi_B)$$

but when it finances other parties it charges them a spread including two CDS spreads

$$(\gamma_L + \pi_L + \pi_B).$$

How to avoid using default as a benefit and at the same time cut down the funding costs so much that banks do not put two CDS in lending spreads? Is maturity transformation the only answer?

When the payoff is wrong

In some cases risk does not reside in unrealistic models, but in the payoff they are applied to, which does not correspond to the real future cashflows of a derivative. We include this case in the model risk topic since these payoff errors are due to misinterpretations similar to those generating model losses. We analyze the recent standard formulas for counterparty risk adjustment and show that they are affected by a strong assumption - probably unnoticed - about the actual payoff at default time (closeout amount). The legal documentation does not seem to underpin this closeout amount. Its adoption, however, has amazing consequences like making a default to infect not only creditors, but also debtors, with a stronger contagion the higher is the default risk of the infected debtor.

CVA and DVA

We consider two parties of a derivative: A (investor) and B (counterparty). We define $\Pi_A(t,T)$ to be the discounted cashflows of the derivative from t to T seen from the point of view of A, with $\Pi_B(t,T) = -\Pi_A(t,T)$. The net present value of the derivative at t is

$$V_A^0(t) := \mathbb{E}_t[\Pi_A(t,T)].$$

The subscript A indicates that this value is seen from the point of view of A, the superscript 0 indicates that we are considering both parties as default-free. Obviously,

$$V_B^0(t) = -V_A^0(t).$$

Unilateral Counterparty Risk

The early literature on counterparty risk adjustment, introduced 'unilateral risk of default': only the counterparty B can default, while the investor A is default free:

$$\begin{aligned} V_A^B(t) &= \mathbb{E}_t\left\{1_{\{\tau^B>T\}}\Pi_A(t,T)\right\} + \\ &+ \mathbb{E}_t\left\{1_{\{\tau^B\le T\}}\left[\Pi_A\left(t,\tau^B\right)+D\left(t,\tau^B\right)\left(R^B\left(V_A^0\left(\tau^B\right)\right)^+-\left(-V_A^0\left(\tau^B\right)\right)^+\right)\right]\right\} \\ &= V_A^0(t)-\mathbb{E}_t\left[L^B 1_{\{\tau^B\le T\}}D\left(t,\tau^B\right)\left(V_A^0\left(\tau^B\right)\right)^+\right] \quad =: \; V_A^0(t)-CVA_A(t)\,. \end{aligned}$$

The approach is easily extended to the case when B is default-free, but the investor A can default:

$$\begin{aligned} V_A^A(t) &= \qquad (62) \\ V_A^0(t)+\mathbb{E}_t&\left[L^A 1_{\{\tau^A\le T\}}D\left(t,\tau^A\right)\cdot\left(-V_A^0\left(\tau^A\right)\right)^+\right] \quad =: \; V_A^0(t)+DVA_A(t)\,. \end{aligned}$$

Bilateral Counterparty Risk

The extension to when both A and B can default is introduced in Picoult (2005), Gregory (2008), Brigo and Capponi (2008), Brigo and Pallavicini (2009). In these previous works the net present value adjusted by the default probabilities of both parties is given by

$$\begin{aligned} V_A(t) &= \mathbb{E}_t\left\{1_0 \Pi_A(t,T)\right\} \\ &+\mathbb{E}_t\left\{1_A\left[\Pi_A\left(t,\tau^A\right)+D\left(t,\tau^A\right)\left(\left(V_A^0\left(\tau^A\right)\right)^+-R^A\left(-V_A^0\left(\tau^A\right)\right)^+\right)\right]\right\} \\ &+\mathbb{E}_t\left\{1_B\left[\Pi_A\left(t,\tau^B\right)+D\left(t,\tau^B\right)\left(R^B\left(V_A^0\left(\tau^B\right)\right)^+-\left(-V_A^0\left(\tau^B\right)\right)^+\right)\right]\right\}, \end{aligned} \tag{63}$$

where we use the following event indicators

$$\begin{aligned} 1_0 &= 1_{\left\{T<\min\left(\tau^A,\tau^B\right)\right\}} \\ 1_A &= 1_{\left\{\tau^A\leq\min\left(T,\tau^B\right)\right\}} \\ 1_B &= 1-1_A-1_0 = 1_{\left\{\tau^B<\tau^A\right\}}1_{\left\{\tau^B\leq T\right\}}. \end{aligned}$$

What happens at default?

We define 1 to be the first entity to default, and 2 to be the second one so that

$$\tau^1 = \min\left(\tau^A, \tau^B\right), \quad \tau^2 = \max\left(\tau^A, \tau^B\right).$$

With these definitions, and recalling that $V_B^0(t) = -V_A^0(t)$, the pricing formula (63) simplifies to

$$\begin{aligned} V_A(t) &= \mathbb{E}_t\left\{1_0 \Pi_A(t,T)\right\} \qquad (64) \\ &+\mathbb{E}_t\left\{(1_B - 1_A)\left[\Pi_2\left(t,\tau^1\right) + D\left(t,\tau^1\right)\left(R^1\left(V_2^0\left(\tau^1\right)\right)^+ - \left(V_1^0\left(\tau^1\right)\right)^+\right)\right]\right\}. \end{aligned}$$

Let us compare this bilateral pricing formula with the price $V_A^B(t)$

$$\begin{aligned} V_A^B(t) &= \mathbb{E}_t\left\{1_{\{\tau > T\}} \Pi_A(t,T)\right\} \qquad (65) \\ &+\mathbb{E}_t\left\{1_{\{\tau \le T\}}\left[\Pi_A(t,\tau) + D(t,\tau)\left(R^1\left(V_A^0(\tau)\right)^+ - \left(V_B^0(\tau)\right)^+\right)\right]\right\}. \end{aligned}$$

What happens at default?

There are precise assumptions on what happens when there is a default event. Start from the unilateral case (65). At τ the residual deal is marked-to-market. The mark-to-market is called *closeout amount*. Here it is given by

$$\left(R^1 \left(V_A^0 (\tau) \right) - \left(V_B^0 (\tau) \right)^+ \right). \tag{66}$$

As indicated by the superscript 0, here the closeout amount $V_A^0 (\tau)$ is computed treating the residual deal as a *default-free* deal. The reason for that is obvious. Now let us look at the pricing formula (64) for bilateral risk of default. The payout at default is now given by

$$(1_B - 1_A) \left[R^1 \left(V_2^0 \left(\tau^1 \right) \right)^+ - \left(V_1^0 \left(\tau^1 \right) \right)^+ \right]$$

Also with bilateral risk of default the closeout amount *is computed treating the residual deal as default-free*. Is this assumption as obviously justified here as it was in the unilateral case? Not quite. Only the assumption of ignoring the risk of default of 1 is justified obviously. But the other party 2 has not defaulted, and there is a non-negligible probability that it defaults before the maturity T of the residual deal.

What happens at default?

If the survived party 2 wanted to substitute the defaulted deal with another one where the market counterparty is default free, the counterparty would ask 2 to pay

$$V_2^2\left(\tau^1\right)=-V_1^2\left(\tau^1\right), \tag{67}$$

because the market counterparty cannot ignore the default risk of party 2 from τ^1 to the maturity T of the residual deal. Thus the amount (67) will be called in the following *substitution closeout amount*. In this case we have

$$\begin{aligned}\hat{V}_A(t) &= \mathbb{E}_t\left\{1_0 \Pi_A(t,T)\right\} \\ &+\mathbb{E}_t\left\{1_A\left[\Pi_A\left(t,\tau^A\right)+D\left(t,\tau^A\right)\left(\left(V_A^B\left(\tau^A\right)\right)^+-R^A\left(-V_A^B\left(\tau^A\right)\right)^+\right)\right]\right\} \\ &+\mathbb{E}_t\left\{1_B\left[\Pi_A\left(t,\tau^B\right)+D\left(t,\tau^B\right)\left(R^B\left(V_A^A\left(\tau^B\right)\right)^+-\left(-V_A^A\left(\tau^B\right)\right)^+\right)\right]\right\}.\end{aligned} \tag{68}$$

ISDA docs

ISDA (2009) Close-out Amount Protocol: "In determining a Close-out Amount, the Determining Party may consider any relevant information, including, without limitation, one or more of the following types of information: (i) quotations (either firm or indicative) for replacement transactions supplied by one or more third parties that may take into account the creditworthiness of the Determining Party at the time the quotation is provided". This is in contrast with the default-free closeout amount $V_2^0\left(\tau^1\right) = -V_1^0\left(\tau^1\right)$ prescribed by the classic formula, and seems instead consistent with the substitution-cost closeout amount $V_2^2\left(\tau^1\right) = -V_1^2\left(\tau^1\right)$ prescribed by the formula given in this paper, that includes the risk of default of the survived party 2. What they imply, in practice?

A practical example

Let us evaluate a 'derivative bond' deal under the assumption that only the counterparty B, namely the borrower, can default. We apply (60) to this payoff, getting

$$V_A^B(0) = e^{-\int_0^T r(s)ds} \Pr\left(\tau^B > T\right) + R^B e^{-\int_0^T r(s)ds} \Pr\left(\tau^B \leq T\right),$$

By applying (62) to this payoff we can compute easily also $V_A^A(0)$, the value when only the lender A can default. We have

$$V_A^A(0) = e^{-\int_0^T r(s)ds}. \tag{69}$$

A practical example

Now we price the deal considering the default risk of both parties, *and assuming first a substitution closeout*. We apply formula (68) introduced in this paper. We have

$$\begin{aligned}\hat{V}_A(0) &= e^{-\int_0^T r(s)ds}\mathbb{E}_0\left[1_{\{\tau^B>T\}}\right] + R^B e^{-\int_0^T r(s)ds}\mathbb{E}_0\left[1_{\{\tau^B\leq T\}}\right] \qquad (70)\\ &= e^{-\int_0^T r(s)ds}\left[\Pr_0\left(\tau^B>T\right) + R^B\Pr_0\left(\tau^B\leq T\right)\right].\end{aligned}$$

The risk of default of the bond-holder in a bond, or the lender in a loan, does not influence the value of the contract. When the contract has no future obligations for a party A the risk of default of A does not influence the price. Only the risk of default of the bond-issuer or borrowers matters. Also this result is in line with market practice.

A practical example

Now we apply instead *the formula (63) that assumes a risk-free closeout*. Since $V_A^0(\tau) = e^{-\int_\tau^T r(s)ds}$, we have

$$V_A(0) \;=\; = e^{-\int_0^T r(s)ds}\mathbb{E}_0[1_0 + 1_A] + e^{-\int_0^T r(s)ds} R^B \mathbb{E}_0[1_B]\,.$$

We can write

$$\begin{aligned} V_A(0) \;&=\; e^{-\int_0^T r(s)ds} \Pr_0\left[T < \min\left(\tau^A, \tau^B\right) \cup \tau^A \le \min\left(T, \tau^B\right)\right] \\ &+ e^{-\int_0^T r(s)ds} R^B \Pr_0\left[\tau^B < \tau^A \cap \tau^B \le T\right]. \end{aligned}$$

and we have

$$V_A(0) \ge \hat{V}_A(0) = V_A^B(0)\,.$$

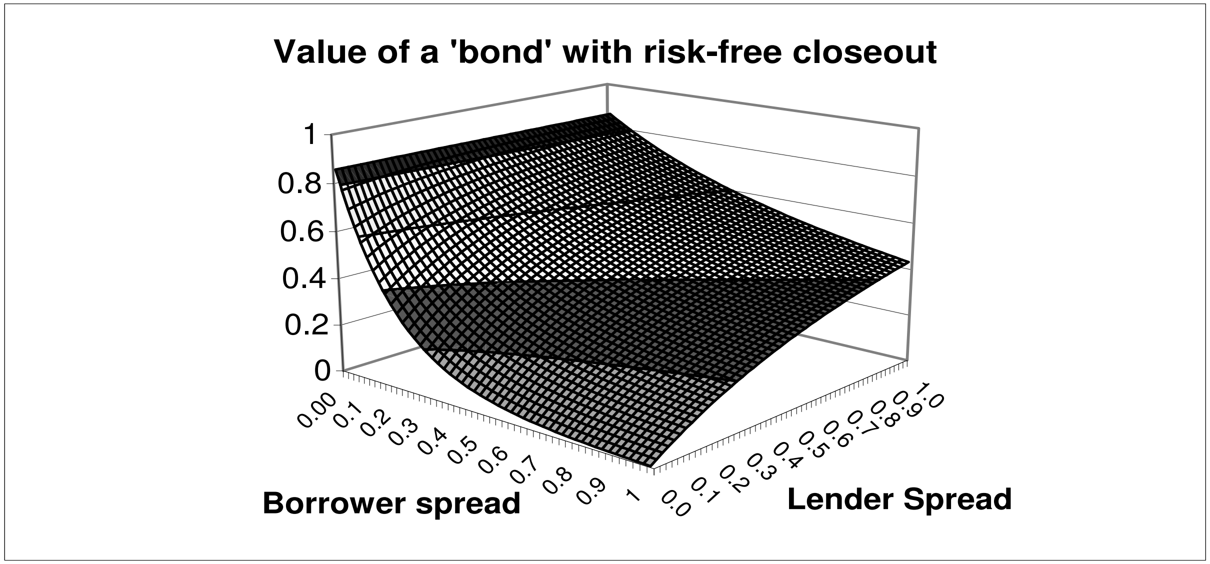
A practical example

Thus a risk-free liquidation increases the value of a 'derivative bond' to the bond holder compared to the value that a bond has in the market practice. Symmetrically, the value is reduced to the bond issuer, and this reduction is an increasing function of the default risk of the bond holder. To quantify the size of the above difference, we take a flat default intensity λ_X. For a cleaner initial analysis, we take τ^A and τ^B to be independent. It is well known that in this case $\min\left(\tau^A, \tau^B\right)$ is also exponentially distributed of parameter $\lambda_A + \lambda_B$, and

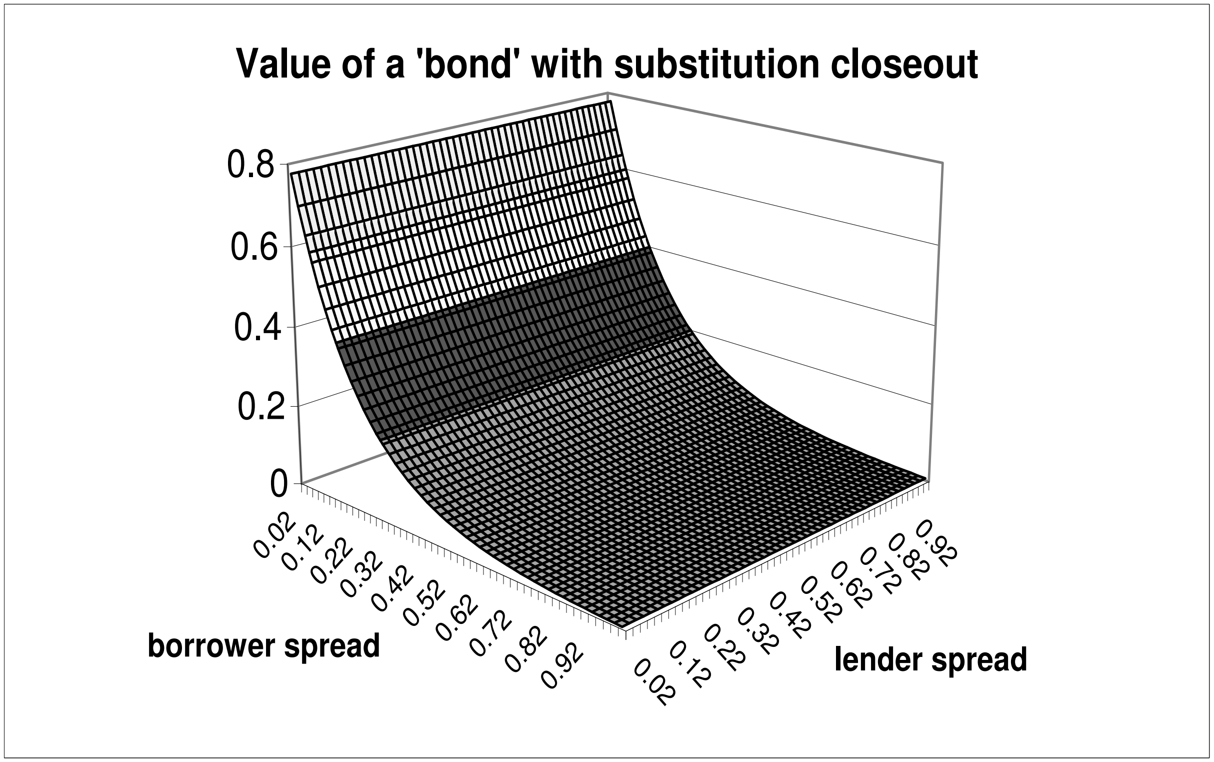
$$\Pr\left(\tau^A < \tau^B\right) = \frac{\lambda_A}{\lambda_A + \lambda_B},$$

We have all we need. We set the risk-free rate at $r = 3\%$, and consider a bond with maturity 5 years, $R^B = 0$. We show the price of the bond for intensities λ_{Lender}, $\lambda_{Borrower}$ going from zero to 100%.

A practical example



A practical example



A practical example

Let us observe a numerical example, for a 1bn notional bond. We suppose the borrower B has a very low credit quality, as expressed by $\lambda_B = 0.2$. The lender A has a much higher credit quality, as expressed by $\lambda_A = 0.04$. A risk free bond issuance with the same maturity and notional would cost $P_T = 860.7mn$. Using the formula with risk-free closeout

$$V_A = 359.5mn$$

to be compared with the price coming form the formula with substitution closeout

$$\hat{V}_A = 316.6mn.$$

The higher value of V_A depends on the probability of default of the lender.

A practical example

With a non-negligible probability, 12%, the lender can default first. Suppose the exact day when default happens is

$$\tau^{Lender} = 2.5 years.$$

Just before default, at 2.5 years less one day, we have for the borrower the following book value of the liability produced by the above deal:

$$V_{Borrower}\left(\tau^{Lender} - 1d\right) = -578.9\text{mln}$$

if he assumes a risk-free closeout, or

$$\hat{V}_{Borrower}\left(\tau^{Lender} - 1d\right) = -562.7\text{mln}$$

if he assumes a substitution closeout.

Contagion to borrowers??

Now default of the lender happens. In case of a risk-free closeout, the book value of the bond becomes simply the value of a risk free bond,

$$V_B\left(\tau^A + 1d\right) = -927.7\text{mln}.$$

The borrower, which has not defaulted, must pay this amount entirely to the defaulted lender. He has a sudden loss of

$$\begin{aligned} & 927.7\text{mln} - 578.9\text{mln} \\ = \quad & 348.8\text{mln} \end{aligned}$$

due to default of the lender. Are we sure this makes sense??