the mortgagor is unable or fails to make future mortgage payments is *default risk*. This risk is mitigated by the house being pledged as collateral for the mortgage, but this will in turn create other risks as outlined below.

- The risk that the value of the property in question falls below the outstanding value of the loan or mortgage. This is often known as the situation of "negative equity" and corresponds to *market risk*. Note that this depends on both the value of the property (collateral) and the value of the mortgage (exposure).
- The risk that the mortgage giver is unable, or faces legal obstacles, to take ownership of the property in the event of the failure to make mortgage payments and faces costs in order to evict the owners and sell the property. This corresponds to *operational or legal risk*.
- The risk that the property cannot be sold immediately in the open market and will have a falling value if property values are in decline. To achieve a sale, the property may then have to be sold at a discount to its fair value if there is a shortage of buyers. This is *liquidity risk*.
- The risk that there is a strong dependence between the value of the property and the default of the mortgagee. For example, in an economic downturn, high unemployment and falling property prices make this rather likely. This is a form of *correlation* (or even *wrong-way*) risk.

The Basics of Collateralisation

The basic idea of collateralisation is very simple and illustrated in Figure 11-1. In a transaction between parties A and B, party A makes a mark-to-market (MtM) profit whilst party B makes a corresponding MtM loss. Party B then posts some form of collateral to party A to mitigate the credit exposure that arises due to the positive MtM. The collateral may be in cash or other securities, the characteristics of which have been agreed before the initiation of the contract.

Note that, since collateral agreements are often bilateral, collateral must be returned or posted in the opposite direction when exposure decreases. Hence, in the case of a positive MtM, an institution will call for collateral and in the case of a negative MtM they will be required to post collateral themselves.

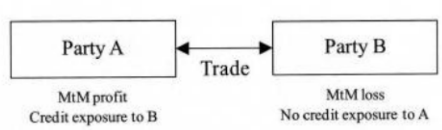


FIGURE 11-1 Illustration of the basic principle of collateralisation.

Collateral Usage

Collateral posting across the market is quite mixed depending on the type of institution (Table 11-1). The main reasons for doing this are the need to post cash or high-quality securities and the operational workload associated with posting collateral under stringent collateral agreements. Other aspects may include external restrictions (e.g., negative pledge provisions) and the economic view that uncollateralised trading is cheaper than collateralised trading. This last point may be considered to be related to CVA charges being too low or not even present.

Nevertheless, collateral usage has increased significantly over the last decade, as illustrated in Figure 11-2, which shows the estimated amount of collateral and gross credit exposure. The ratio of these quantities gives an estimate of the fraction of credit exposure that is collateralised. This has grown year on year to a ratio of around 90% in 2009 but is a slightly misleading figure due to the presence of overcollateralisation. Nevertheless, the impact of collateralisation is reported to reduce overall exposure by around four-fifths (Ghosh et. al., 2008). Incorporating the fact that credit exposures are first decreased through netting and the remaining net exposures are further mitigated by the pledging of collateral reduces total market exposure by nearly 93% (Bliss and Kaufman, 2005).

TABLE 11-1 Collateral Posting by Type of Institution

Institution Type	Collateral Posting
Dealer Banks	Very High
Other Banks	High
Supranationals, Local Authorities, Private Equity Funds	Low
Corporates	Low
Sovereigns	Very Low

Source: Adapted from ISDA Market Review of OTC Derivatives Bilateral Collateralisation Practices 36–38 (2010).

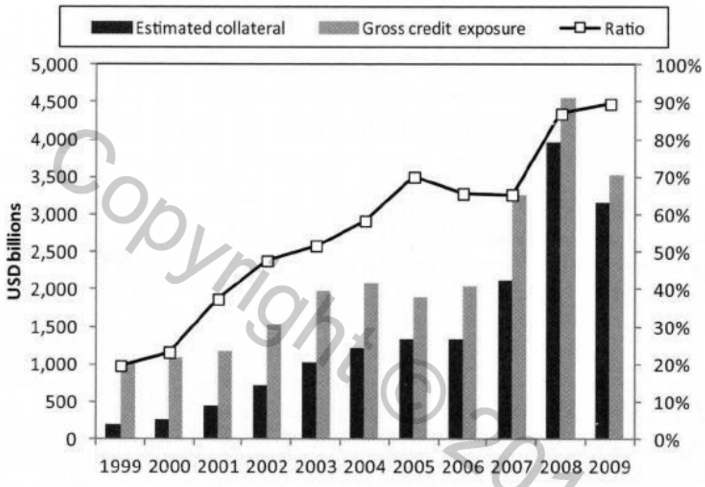


FIGURE 11-2 Illustration of the amount of collateral compared to the gross credit exposure (i.e., after netting has been accounted for) and the ratio given the extent of collateralisation of OTC derivatives.

Source: ISDA and BIS.

The Credit Support Annex

Within an ISDA Master Agreement (see previous chapter), it is possible to append a credit support annex (CSA) which permits the parties to mitigate their credit risk further by agreeing to various collateral posting.[1] As with netting, ISDA has legal opinions throughout a large number of jurisdictions regarding the enforceability of the provisions within a CSA. The CSA is therefore at the centre of any collateral agreement as it governs the mechanics of collateral with respect to issues such as:

- Method and timings of the underlying valuations.
- The calculation of the amount of collateral that will be posted.
- The mechanics and timing of collateral transfers.
- Eligible collateral.
- Collateral substitutions.
- Dispute resolution.
- Interest rate payments on collateral.
- Haircuts applied to collateral securities.
- Possible rehypothecation (reuse) of collateral securities.
- Triggers that may change the collateral conditions (for example, ratings downgrades that may lead to enhanced collateral requirements).

In addition, the nature of a CSA is critically defined by a number of key parameters that essentially define the amount of collateral that will be posted. The most important parameters are:

- *Threshold*. Defines the level of MtM above which collateral is posted. When the exposure is above the threshold, the threshold amount is *undercollateralised*. When the exposure is below the threshold then it is not collateralised at all.
- *Independent amount*. This defines an amount of extra collateral that must be posted irrespective of the exposure. Hence, the exposure is *overcollateralised*. An independent amount is similar in concept to an initial margin required by an exchange or central counterparty. An independent amount is not common in CSAs, although it is used in some specific cases and will be required under new regulations in many cases (e.g., inter-bank trades).
- *Minimum transfer amount*. This defines the minimum amount of collateral that can be called for at a time.

Note that thresholds and independent amounts essentially work in opposite directions. Mathematically, an independent amount is a negative threshold and vice versa.

The process by which two counterparties will agree to collateralise their exposures can be summarised as follows:

- Parties negotiate and sign a collateral support document containing the terms and conditions under which they will operate.
- Trades subject to collateral are regularly marked-to-market, and the overall valuation including netting is agreed (unless this amount is disputed as discussed later).
- The party with negative MtM delivers collateral (subject to minimum transfer amounts and thresholds as discussed later).
- The collateral position is updated to reflect the transfer of cash or securities.
- (Periodic reconciliations should also be performed to reduce the risk of disputes.)

CSAs must explicitly define all the parameters of the collateralisation and account for all possible scenarios. The choice of parameters will often come down to a balance between the workload of calling and returning collateral

[1] 92% of collateral agreements in use are ISDA agreements. *Source:* ISDA Margin Survey 2010.

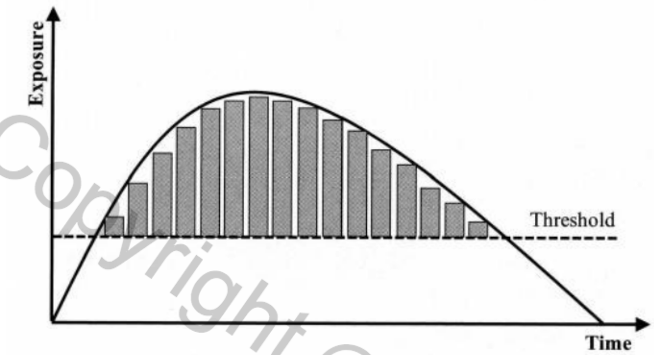


FIGURE 11-3 Illustration of the impact of collateral on exposure. The collateral amount is depicted by the grey areas.

versus the risk mitigation benefit of doing so. We will now analyse the components that make up the collateral process in more detail.

Impact of Collateral

The impact of collateral on a typical exposure profile is shown in Figure 11-3. There are essentially two reasons why collateral cannot perfectly mitigate exposure. Firstly, the presence of a threshold[2] means that a certain amount of exposure cannot be collateralised. Secondly, the delay in receiving collateral and parameters such as the minimum transfer amount create a discrete effect, as the movement of exposure cannot be tracked perfectly.[3] This is illustrated by the grey blocks in Figure 11-3.

COLLATERAL TERMS

Valuation Agent

The valuation agent is normally the party calling for delivery or return of collateral and thus must handle all calculations. Large counterparties trading with smaller counterparties may insist on being valuation agents for all purposes. In such a case, the "smaller" counterparty is not obligated to return or post collateral if they do not receive the expected notification, whilst the valuation agent is under obligation to make returns where relevant. Alternatively, both counterparties may be the valuation agent and each will call for collateral when they have an exposure.

The role of the valuation agent in a collateral calculation is as follows:

- To calculate credit exposure under the impact of netting.
- To calculate the market value of collateral previously posted.
- To calculate the uncollateralised exposure.
- To calculate the delivery or return amount (the amount of collateral to be posted by either counterparty). This is likely to differ from the uncollateralised exposure due to the discrete nature of collateral agreements, which means that collateral is transferred in blocks.

Third-party valuation agents provide operational efficiencies, and can also help prevent disputes that are common in bilateral collateral relationships.

Types of Collateral

Cash is the major form of collateral taken against OTC derivatives exposures (Figure 11-4). The ability to post other forms of collateral is often highly preferable for liquidity reasons, but the credit crisis has shown that even government agency securities (for example, Fannie Mae and Freddie Mac) and Triple-A MBS securities are far from the high-quality assets with minimal price volatility that they were once assumed to be. Non-cash collateral also creates the problems of reuse or rehypothecation (discussed later) and additional volatility arising from the price uncertainty of collateral posted and its correlation to the original exposure. On the contrary, in extreme market conditions, cash tends to be in limited supply.

If the credit rating of an underlying security held as collateral declines below that specified in the collateral agreement, then normally it will be necessary to replace this security immediately. When two counterparties do not have the same local currency, one of them will have to take FX risk linked to the collateral posted, even when it is in the form of cash. Securities in various currencies may be specified as admissible collateral but may also attract larger haircuts due to the additional FX risk. FX risk from posted collateral can be hedged in the spot and forward FX markets but it must be done dynamically as the value of collateral changes.

[2] Note that a threshold can be zero, in which case this is not an issue. However, even many interbank CSAs have non-zero thresholds.

[3] The purpose of an independent amount is to mitigate this risk by providing a buffer.

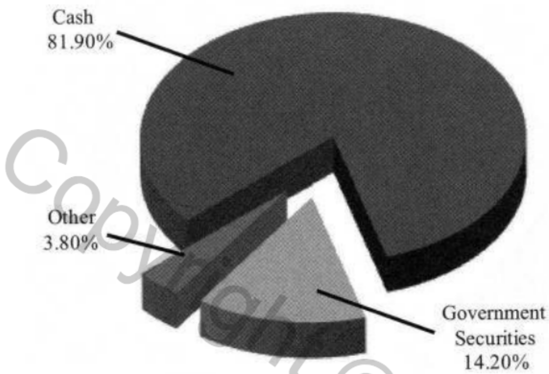


FIGURE 11-4 Breakdown of the type of collateral used for OTC derivatives.

Around three-quarters of cash is posted in USD and EUR currencies with more than 97% being covered by these two currencies plus GBP and JPY. Likewise, over 95% of government securities are posted in these four currencies. The other category contains collateral such as government agency securities, supranational bonds, covered bonds, corporate bonds, letters of credit, and equities.

Source: ISDA Margin Survey 2010.

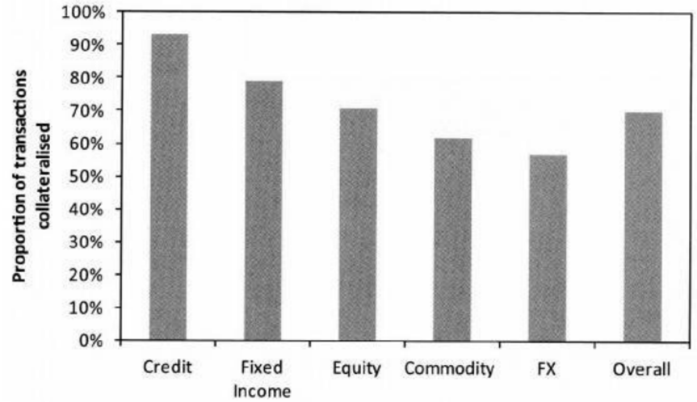


FIGURE 11-5 Illustration of the proportion of OTC derivatives collateralised shown by product type.

Source: ISDA Margin Survey 2010.

Coverage of Collateralisation

As illustrated in Figure 11-5, a large proportion of all OTC derivatives trade under collateral agreements. The percentages are highest for credit derivatives, which is not surprising due to the high volatility of credit spreads[4] whilst the fact that many FX transactions are short-dated explains the relatively low number for this asset class.

Collateral agreements will reference the netted value of some or all trades with a specific counterparty. From a risk mitigation point of view, one should include the maximum number of trades but this should be balanced against the need to effectively value all such trades. Product and regional impacts are often considered when excluding certain trades from collateral agreements. Collateral agreements do require the transfer of the undisputed amount immediately, which means that the majority of products should still be collateralised even when there are disputes regarding a minority. However, the cleaner approach of leaving such products outside a collateral agreement is sometimes favoured.

[4] In addition, the wrong-way risk embedded in credit derivatives may be driving this aspect although it is not clear whether collateral strongly mitigates this component.

Disputes and Reconciliations

Collateral management is one of the few areas of banking that has not fully embraced technological advances and still relies heavily on manual process and data standards. The use of spreadsheets is still quite common. Clearly, such practices can lead to significant disputes between counterparties.

A dispute over a collateral call is common and can arise due to one or more of a number of factors:

- trade population;
- trade valuation methodology;
- application of netting rules;
- market data and market close time;
- valuation of previously posted collateral.

If the difference in valuation or disputed amount is within a certain tolerance specified in the collateral agreement, then the counterparties may "split the difference." Otherwise, it will be necessary to find the cause of the discrepancy. Obviously, such a situation is not ideal and will mean that one party will have a partially uncollateralised exposure at least until the origin of the disputed amount can be traced, agreed upon and corrected. The following steps are normally followed in the case of a dispute:

- The disputing party is required to notify its counterparty (or the third-party valuation agent) that it wishes to dispute the exposure or collateral calculation no later than the close of business on the day following the collateral call.

- The disputing party agrees to transfer the undisputed amount and the parties will attempt to resolve the dispute within a certain timeframe (the "resolution time"). The reason for the dispute will be identified (e.g., which transactions have material differences in valuation).
- If the parties fail to resolve the dispute within the resolution time, they will obtain MtM quotations from several market makers for the components of the disputed exposure (or value of existing collateral in case this is the component under dispute).

Rather than being *reactive* and focussing on dispute resolution, it is better to be *proactive* and aim to prevent disputes in the first place. Reconciliations aim to minimise the chance of a dispute by agreeing on valuation figures even though the resulting netted exposure may not lead to any collateral changing hands. They can even be performed using dummy trades before two counterparties transact with one another. It is good practice to perform reconciliations at periodic intervals (for example, weekly or monthly) so as to minimise differences in valuation between counterparties. Such reconciliations can pre-empt later problems that might arise during more sensitive periods. Reconciliations may be rather detailed and will therefore highlight differences that otherwise may be within the dispute tolerance or that by chance offset one another. Hence, problems that may otherwise appear only transiently should be captured in a thorough reconciliation.

The global financial crisis highlighted many problems in the collateral management practices of banks. Regulators have reacted to this in the Basel III proposals, which give less credit to collateral as a risk mitigant in some cases. Collateral management is improving, mainly via a simplification of collateral terms (e.g., using cash collateral only). This is also driven by the issues that collateral creates in terms of valuation.

Margin Call Frequency

Margin call frequency refers to the periodic timescale with which collateral may be called and returned. A longer margin call frequency may be agreed upon, most probably to reduce operational workload and in order for the relevant valuations to be carried out. Some smaller institutions may struggle with the operational and funding requirements in relation to the daily margin calls required by larger counterparties. Whilst a margin call frequency longer than daily might be practical for asset classes and markets that are not so volatile, daily margining is becoming standard in OTC derivatives markets. Furthermore, intraday margining is common for vanilla products such as repos and for derivatives cleared via central counterparties.

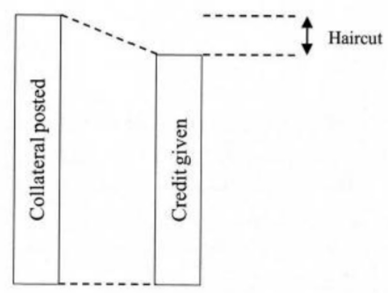


FIGURE 11-6 Illustration of a haircut applied to collateral.

Haircuts

A haircut is a discount applied to the value of collateral to account for the fact that its value may deteriorate over time. Cash collateral in a major currency will require no haircut but other securities will have pre-specified haircuts depending on their individual characteristics. A haircut of x% means that for every unit of that security posted as collateral, only $(1 - x)$% of credit (or "valuation percentage") will be given, as illustrated in Figure 11-6. The collateral giver must account for the haircut when posting collateral.

Haircuts are primarily used to account for the price volatility of collateral posted. Default relating to a security posted as collateral would clearly reduce the value of that collateral substantially and the haircut is very unlikely to cover such an event. For this reason, only high-quality debt securities are typically allowed to be used as collateral. Haircuts are designed to cover price volatility of assets only and it is therefore crucial that every effort is made to avoid collateral with significant default risk.

Some examples of haircuts together with eligible collateral types are shown in Table 11-2.

The important points to consider before assigning a haircut are:

- time taken to liquidate the collateral;
- volatility of the underlying market variable(s) defining the value of the collateral;
- default risk of the security;
- maturity of the security;

TABLE 11-2 Example Haircuts in a Collateral Agreement

	Party A	Party B	Valuation Percentage	Haircut
Cash in eligible currency	X	X	100%	0%
Debt obligations issued by the governments of the US, UK or Germany with a maturity less than 1 year	X	X	98%	2%
Debt obligations issued by the governments of the US, UK or Germany with a maturity between 1 and 10 years	X	X	95%	5%
Debt obligations issued by the governments of the US, UK or Germany with a maturity greater than 10 years	X		90%	10%

- liquidity of the security;
- any relationship between the default of the counterparty and the value of the collateral (wrong-way risk).

For example, a high-quality long-dated government or corporate bond has significant interest rate volatility due to the long maturity, although default and liquidity risk will probably not be of great concern. Such a security might therefore attract a haircut of around a few percent. Collateral with greater credit risk must be assigned a larger haircut to account for the credit spread volatility and default risk. Volatile collateral such as commodities (e.g., gold) and equities should be assigned higher haircuts as compensation for their price volatility and potential illiquidity. Finally, it is important to consider the potential correlation between the exposure and the valuation of collateral.

Example

Consider a security that attracts a haircut of 5% and is being posted to cover a collateral call of $100,000. Only 95% of the value of this security is credited for collateral purposes and so the actual amount of collateral posted must be

Market value of collateral	= $105,263
Haircut	= $5,263 (5% of $105,263)
Credit given	= $100,000 (difference between the above)

It is the collateral giver's responsibility to account for haircuts when posting collateral so that if a collateral call is made as above, then (assuming they do not dispute the amount) the counterparty could post $100,000 in cash but $105,263 in terms of the market value of a security attracting a 5% haircut.

Coupons and Interest Payments

As long as the giver of collateral is not in default, then they remain the owner from an economic point of view. Hence, the receiver of collateral must pass on coupon payments, dividends and any other cash flows. The only exception to this rule is in the case where an immediate margin call would be triggered. In this case, the collateral receiver may typically keep the minimum component of the cash flow in order to remain appropriately collateralised.

Interest will typically be paid on cash collateral at the overnight indexed swap (OIS) rate (for example, EONIA in Europe, Fed Funds in the U.S.). The logic behind this is that since collateral may only be held for short periods (due to the variation of exposure), only a short-term interest rate can be paid. However, OIS is not necessarily the most appropriate collateral rate, especially for long-dated exposures where collateral may be held in substantial amounts for a long period. This may lead to a negative carry problem due to an institution funding the collateral posted at a rate significantly above LIBOR but receiving only the OIS rate (less than LIBOR) for the collateral posted. Sometimes, a collateral receiver may possibly agree to post a rate higher than OIS to compensate for this funding mismatch. Another reason for a collateral receiver to pay a return in excess of OIS would be to incentivise the posting of cash over other more risky and volatile securities.

Substitution, Funding Costs and Rehypothecation

Sometimes a counterparty may require or want securities posted as collateral returned (for example, to meet delivery commitments).[5] In this case, they can make a substitution request and post an alternative amount of eligible collateral with the relevant haircut applied. The requested collateral does not need to be released until the substitute collateral has been received. If substitution is optional (no consent required), then a substitution request cannot be refused[6] (unless it is not valid). Alternatively, substitution may only be allowed if the holder of the collateral gives consent. Whether or not collateral can be substituted without consent is an important consideration in terms of the funding costs and benefits of collateral.

For collateral to provide benefit against funding costs, it must be usable (since the economic ownership remains with the collateral giver) via rehypothecation, which means it can be posted as collateral or pledged via repo. To understand the importance of this, consider Figure 11-7. Collateral in securities that cannot be rehypothecated reduces counterparty risk but creates a funding problem. We will refer to this as funding liquidity risk.

Rehypothecation would seem to be obvious because it keeps the flow of collateral moving around the financial system without any blockage. The question arises as to whether rehypothecating a security in this way creates additional risk due to a loss of control of collateral. An institution faces two possible risks in this respect:

- Collateral pledged in a collateral agreement against a negative MtM to another counterparty may be rehypothecated and consequently not be returned (in the event of a default of the counterparty coupled to an increase in the MtM).
- Collateral received from party A and then rehypothecated to party B. This may not be retrieved in the event that party B defaults, creating a liability to party A.

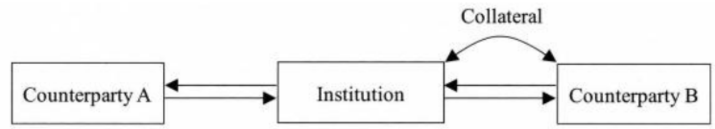


FIGURE 11-7 Illustration of the importance of rehypothecation of non-cash collateral.

An institution trades with counterparty A and typically hedges this transaction with counterparty B, both under CSA agreements. If counterparty B posts collateral then ideally it should be possible to pass this collateral on to counterparty A to minimise funding costs.

Prior to the credit crisis in 2007, the pledging, reuse and rehypothecation of collateral was strongly encouraged. This was viewed as being critical to the entire financial system (Segoviano and Singh, 2008). However, the practice of rehypothecation probably became too widespread, especially in the interbank market (presumably since there was little concern of actual bank defaults). The bankruptcy of Lehman Brothers has illustrated the potential problems with rehypothecation. One example is that customers of Lehman Brothers Inc. (US) were being treated more favourably than the UK customers of Lehman Brothers International (Europe) in terms of the return of rehypothecated assets (due to differences in customer protection between the UK and the US[7]).

Singh and Aitken (2009) have reported a significant drop in rehypothecation, which is safer from a systemic risk perspective but leads to an increase in funding liquidity risk. Hedge funds are tending to be unwilling to allow rehypothecation, which will surely lead to an increase in prime broker fees. The problems with rehypothecation are another driving force behind cash collateralisation becoming increasingly the standard and, in many cases, the only option that most institutions are willing to adopt.

When posting and receiving collateral, institutions are becoming increasingly aware of the need to optimise their collateral management as, during the financial crisis, funding efficiencies have emerged as an important driver of collateral usage. Collateral management is no

[5] Note that the collateral returned needs not be exactly the same but must be equivalent (e.g., the same bond issue).

[6] For example, on grounds that the original collateral has been repoed, posted to another counterparty, sold or is otherwise inaccessible.

[7] The liquidator of Lehman Brothers (PWC) stated in October 2008, shortly after the bankruptcy, that certain assets provided to Lehman Brothers International (Europe) had been rehypothecated and may not be returned.

longer a back-office cost centre but can be an important asset optimisation tool delivering the most cost-effective collateral. An institution must consider the "cheapest-to-deliver" cash collateral and account for the impact of haircuts and the ability to rehypothecate non-cash collateral. For example, different currencies of cash will pay different OIS rates and non-cash collateral, if rehypothecated, will earn different rates on repo.

DEFINING THE AMOUNT OF COLLATERAL

Types of CSA

Due to the very different nature of OTC derivatives counterparties, many different collateral arrangements exist. Broadly speaking, these can be categorised into the following.

(i) No CSA

There are two reasons why an institution may be unable or unwilling to post collateral. Firstly, it could be because their credit quality is far superior to their counterparty. Secondly (or additionally), it may occur because they cannot commit to the operational and liquidity requirements that arise from committing to a CSA.

One result of the above is that in some trading relationships, CSAs are not used because one or both parties cannot commit to collateral posting. A typical example of this is the relationship between a bank and a corporate where the latter's inability to post collateral means that a CSA is not usually in place (for example, a corporate treasurer may find it almost impossible to manage their liquidity needs if they transacted under a CSA).

(ii) Two-way CSA

For two similar counterparties, a two-way CSA is more typical. This is common, for example, in the interbank market. A two-way CSA is typically beneficial to both parties. Two-way CSAs may be skewed in some way. For example, one party may have a lower threshold than the other, which may be due to their inferior credit rating.

(iii) One-way CSA

In some situations, a one-way CSA is used which is beneficial to only the collateral receiver. Indeed, a one-way CSA represents additional risk for the collateral giver that puts them in a worse situation than if they were in a no-CSA relationship. One example of this would be a bank trading with a hedge fund and requiring collateral posting (possibly including an independent amount) to mitigate the significantly increased (and opaque) counterparty risk of the hedge fund. Another typical example is a high-quality entity such as a Triple-A sovereign or insurer trading with a bank.

Note that not all one-way CSAs are truly one-way. For example, one party may not post collateral immediately but may be required to do so if, for example, their credit rating deteriorates. Prior to the financial crisis, Triple-A entities such as monoline insurers traded through one-way CSAs but with triggers specifying that they must post collateral if their ratings were to decline. This seemed to put banks in a safe position but quite the reverse was true.

Linkage of Collateral Parameters to Credit Quality

It is quite common to attempt to link the precise terms of a collateral agreement to the credit quality of one or both counterparties. The motivation for doing this is to minimise the operational workload whilst a counterparty has strong credit quality but have the ability to tighten up the terms of collateralisation when their credit quality deteriorates. The quantities to which collateral terms can obviously be linked are:

- credit ratings;
- traded credit spread;
- market value of equity;
- net asset value (sometimes used in the case of hedge funds).

The most commonly used of the above have been credit ratings, and examples will be given below. Linking a tightening of collateral terms to a credit rating (for example, a downgrade to sub-investment grade) might seem a rather easy and obvious method of mitigating an increase in counterparty risk. However, this type of agreement can lead to rather unpleasant discontinuities since a downgrade of a counterparty's credit rating can occur rather late and then cause further credit issues due to the requirement to post collateral (similar to the discussions around additional termination events).

TABLE 11-3 Illustration of Linkage of Threshold to Credit Rating

Rating	Threshold
AAA	$100m
AA	$50
A	$25m
Lower	Zero

Threshold

A threshold is a level of exposure below which collateral will not be called. The threshold therefore represents an amount of uncollateralised exposure. If the exposure is above the threshold, only the *incremental* exposure will be collateralised. In return for taking the risk of a moderate uncollateralised exposure, the operational burden of calling and returning collateral will be reduced. Put another way, many counterparties may only consider collateralisation important when the exposure exceeds a certain level, the threshold. A threshold of zero implies that any exposure is collateralised, whilst a threshold of infinity is used to specify that a counterparty will not post collateral under any circumstance. An example of thresholds and their linkage to credit rating is shown in Table 11-3.

A downgrade of the counterparty may trigger an immediate collateral call. As discussed previously, if such an agreement is in place with many counterparties then it may cause cash flow issues at precisely the worst time. This is exactly what happened with AIG and mono-line insurers and will be discussed in more detail in Chapter 16.

Independent Amount

An independent amount can be thought of (intuitively and mathematically) as a negative threshold. It is typically held as a cushion against "gap risk," the risk that the market value of a transaction(s) may gap substantially in a short space of time. An independent amount can be significant and reduce exposure to practically zero. Sometimes the posting of an independent amount may be linked to a downgrade in a counterparty's credit rating.

We can think of an independent amount as transforming counterparty risk into "gap risk." A transaction with a risky counterparty might be collateralised with both frequent margin calls and additionally an independent amount. The aim is then that the transaction is always overcollateralised by the independent amount so that even if the counterparty defaults, it is highly unlikely that any loss will be suffered. The residual risk is that, when the counterparty defaults, the value of the transactions will move dramatically or "gap" before it can be unwound. The independent amount is often considered large enough to make such a gap event in the relevant time horizon highly unlikely.

Independent amounts are often specific to a particular trade and are common for counterparties considered to be of relatively poor credit quality (such as hedge funds). However, future regulation seems likely to make them much more common.

Minimum Transfer Amount and Rounding

Collateral cannot be transferred in blocks that are smaller than the minimum transfer amount and hence this must be considered when calculating the amount of collateral that could be called. This will typically mean that an increasing exposure will be slightly under-collateralised due to minimum transfer restrictions. In contrast, a decreasing exposure will typically mean an institution has a small overcollateralisation since they do not need to return collateral continuously.

A minimum transfer amount is the smallest amount of collateral that can be transferred. It is used to avoid the workload associated with a frequent transfer of insignificant amounts of collateral. The size of the minimum transfer amount again represents a balance between risk mitigation versus operational workload. The minimum transfer amount and threshold are additive in the sense that the exposure must exceed the sum of the two before any collateral can be called. We note this additively does not mean that the minimum transfer amount can be incorporated into the threshold—this would be correct in defining the point at which the collateral call can be made but not in terms of the collateral due.[8] Again, minimum

[8] For example, for a minimum transfer amount (MTA) and threshold (K), the collateral call can be made when the (potentially already collateralised) exposure (E) exceeds MTA + K but the collateralisation required is $E - K$. Using the approximation of adding MTA and K would (conservatively) model the collateralisation required as $E - K -$ MTA.

transfer amounts may be linked to ratings. When the counterparty has a weaker credit rating then the additional operational workload required to make a larger number of smaller collateral calls is a reasonable price to pay for being able to reduce the exposure.

A collateral call or return amount will always be rounded to a multiple of a certain size to avoid unnecessarily small amounts. The rounding may be always up (or down) or might always be in favour of one counterparty (i.e., up when they call for collateral and down when they return collateral). This is typically a relatively small amount and will have a small effect on the impact of collateralisation. However, the impact of rounding can be considered alongside the other factors above and will cause minor but noticeable impacts on the overall exposure.

THE RISKS OF COLLATERALISATION

Whilst collateral management is a very useful tool for mitigating counterparty risk, it has significant limitations that must be considered. Essentially, the counterparty risk is converted into other forms of financial risk, such as legal risk (for example, if the terms defined in a CSA cannot be upheld within the relevant jurisdiction). Correlation risk (where collateral is adversely correlated to the underlying exposure), credit risk (where the collateral may default or suffer an adverse credit effect), and FX risk (due to collateral being posted in a different currency) are also important. However, the three most important risks to consider are market, operational and liquidity risk.

Market Risk and the Margin Period of Risk

Collateral can never completely eradicate counterparty risk and we must consider the residual risk that remains under the collateral agreement. This is mainly due to contractual parameters such as thresholds and minimum transfer amounts that effectively delay the collateral process, in addition to the normal delay, since collateral cannot be received immediately. This can be considered a market risk as it is related to the extent of market movements after the counterparty last posted collateral. Whilst the residual risk may be only a fraction of the uncollateralised risk, it may be more difficult to quantify and indeed hedge.

Whilst there is a contractual period between collateral calls (often daily), one must consider what we shall call the "margin period of risk". This is the *effective* time assumed between a collateral call and receiving the appropriate collateral (or in a worst-case scenario putting the counterparty in default, liquidating existing collateral, closing out and re-hedging the trade). Under Basel II regulations, this should be a *minimum* of 10 days for OTC derivatives and the new Basel III regime defines a more conservative 20-day minimum in certain cases. The experience of most participants in the Lehman Brothers bankruptcy was that this period was approximately 5–10 business days.

Operational Risk

The time-consuming and intensely dynamic nature of collateralisation means that operational risk is a very important aspect. The following are examples of specific operational risks:

- missed collateral calls;
- failed deliveries;
- computer error;
- human error;
- fraud.

There is clearly no point in having a collateral management programme that reduces significantly many credit exposures only to find that, in the event of an actual default, losses are not mitigated due to some lack of control or error. The following is a list of points to consider in relation to operational risk:

- Legal agreements must be accurate and enforceable.
- IT systems must be capable of automating the many daily tasks and checks that are required.
- The regular process of calls and returns of collateral is complex and can be extremely time-consuming with a workload that increases in markets that are more volatile.
- Timely accurate valuation of all products is key.
- Information on independent amounts, minimum transfer amounts, rounding, collateral types and currencies must be maintained accurately for each counterparty.
- Failure to deliver collateral is a potentially dangerous signal and must be followed up swiftly.

Liquidity Risk

Collateralisation of counterparty risk leads to demanding liquidity requirements. Indeed, this is why some counterparties do not sign CSAs in the first place. One of the most obvious manifestations of this liquidity risk is in the event that collateral has to be liquidated following the default of a counterparty. Firstly, the surviving institution faces transaction costs (bid–offer) and market volatility over the liquidation period. Secondly, there is the risk that by liquidating an amount of a security that is large compared with the volume traded in that security, the price will be driven down and a potentially large loss incurred. If one chooses to liquidate the position more slowly in small blocks, then there is exposure to market volatility for a longer period.

When agreeing to collateral that may be posted and when receiving securities as collateral, important considerations are:

- What is the total issue size or market capitalisation posted as collateral?
- Is there a link between the collateral value and the credit quality of the counterparty? Such a link may not be obvious and predicted by looking at correlations between variables.[9]
- How is the relative liquidity of the security in question likely to change if the counterparty concerned is in default?

Because of these liquidity impacts, a concentration limit of 5–10% may be imposed to prevent severe liquidation risk in the event of a counterparty defaulting.

Funding Liquidity Risk

The above considerations only come into play when a counterparty has actually defaulted. A more significant aspect of liquidity risk arises from the funding needs that arise due to CSAs. We refer to this as funding liquidity risk.

Despite the increased use of collateral, a significant portion of OTC derivatives remain uncollateralised. This arises mainly due to the nature of the counterparties involved, such as corporates and sovereigns, without the liquidity and operational capacity to adhere to daily collateral calls. In such cases, an institution must consider the funding implications that arise. Since most banks aim to run mainly flat (hedged) OTC derivatives books, funding costs arise from the nature of hedging: a non-CSA trade being hedged via a trade done within a CSA arrangement. This relationship was illustrated in Figure 11-7, where the institution will incur a funding cost when the uncollateralised trade moves in their favour and experience a benefit when the reverse happens. In the recent financial crisis where funding has become costly, the need to assess this carefully has become paramount.

The implications of funding liquidity risk arising from CSAs are even more important for non-banking organisations such as institutional investors, corporates, and sovereigns that trade under CSAs with banks.[10] The conversion of counterparty risk into funding liquidity risk will be beneficial in normal, liquid markets where funding costs are low. However, in abnormal markets where liquidity is poor, funding costs can become significant and may put extreme pressure on an institution.

Below is a good example of the dangers of funding liquidity risk. By not demanding collateral, a bank is essentially converting this risk back into counterparty risk, presumably with the view that long-term counterparty risk is better than short-term liquidity risk. The mitigation of counterparty risk via collateral is clearly useless in such a case.

The BP Deepwater Horizon Oil Spill

In 2010, British Petroleum (BP) experienced the largest accidental marine oil spill in the history of the petroleum industry. This caused loss of life, severe environment problems and, of course, severe financial losses for BP themselves. In the immediate aftermath, some banks gave some flexibility to BP in terms of collateral posting. An obvious way to interpret this is that the banks believed that, whilst BP was certainly experiencing significant idiosyncratic credit problems, it was unlikely to default. Forcing the contractual posting of collateral (which may have been triggered by the resulting credit rating downgrades of the company) may have caused BP liquidity problems that would have made their default more likely.

[9] In the case of the Long-Term Capital Management (LTCM) default, a very large proprietary position on Russian government bonds made these securities far from ideal as collateral. Even a European bank posting cash in Euros gives rise to a potentially problematic linkage.

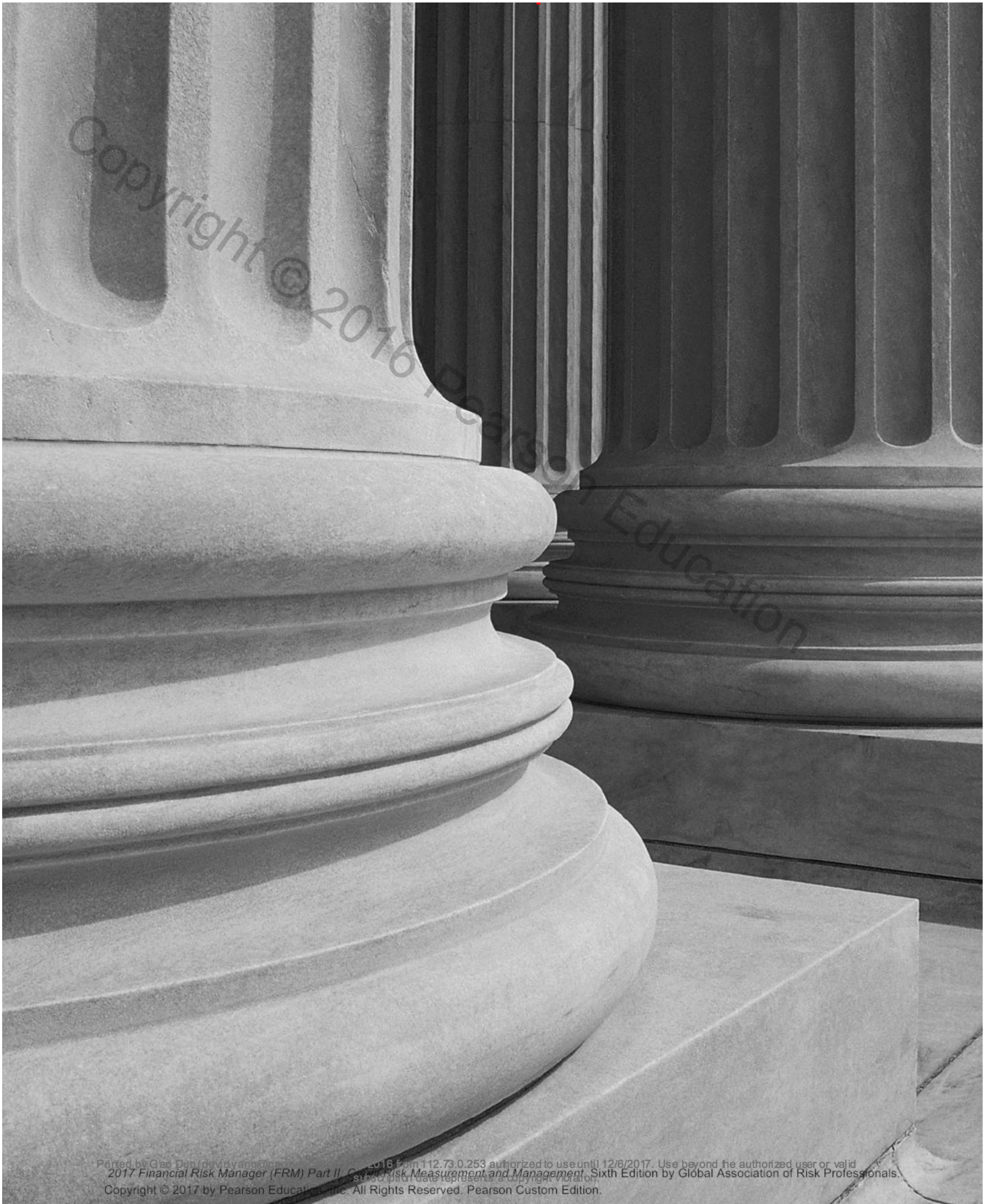
[10] Note that the posting of collateral here by the institutions mentioned may be subject to downgrade triggers.

SUMMARY

In this chapter we have discussed in detail the use of collateral management in controlling credit exposure, which is a crucial method when trading involving large positions and/or relatively risky counterparties. We have described the mechanics of collateral management and the variables that determine how much collateral would be posted. The significant risks that arise from collateral use have also been considered.

Collateral management should be understood as a way to improve recovery in the event of a counterparty actually defaulting but it is certainly not a replacement for a proper ongoing assessment of credit quality and quantification of credit exposure. Furthermore, the use of collateral mitigates counterparty risk but can aggravate funding liquidity risk and create other financial risks.

In our description of risk mitigation we have now covered all the methods of reducing credit exposure.



Central Counterparties 12

Learning Objectives

After completing this reading you should be able to:

- Explain the objectives and functions of central counterparties (CCPs).
- Discuss the strengths and weaknesses of CCPs.
- Describe the different CCP netting schemes, the benefit of netting, and distinguish between bilateral netting and multilateral netting.
- Discuss the key challenges in relation to the clearing of over-the-counter (OTC) derivative products.
- Describe the three types of participants that channel trade through the CCP.
- Explain the loss waterfall in a CCP structure.
- Define initial margin and variation margin, and describe the different approaches and factors in calculating initial margin.
- Discuss the impact of initial margin on prices, volume, volatility, and credit quality.
- Explain factors that can lead to failure of a CCP and discuss measures to protect CCPs from default.

Excerpt is Chapter 7 of Counterparty Credit Risk and Credit Value Adjustment, *Second Edition, by Jon Gregory.*

Put all your eggs in one basket—and watch that basket!

—Mark Twain

CENTRALISED CLEARING

In this chapter. we consider the role of central counterparties (CCPs) to provide a means for centralisation, mutualisation and reduction of counterparty risk. Following the global financial crisis that started in 2007, there has been a significant regulatory interest in expanding the role of CCPs for mitigating counterparty risk. In particular, the interest has been strong for credit derivatives products with their embedded wrong-way risks. Channelling OTC derivatives transactions through CCPs has two main objectives. The first is to reduce counterparty credit risk. The second is to increase transparency so that regulators are more easily able to quantify the positions being taken and carry out stress tests. We will discuss the function of a CCP in detail and aim to highlight the strengths, weaknesses and possible unintended consequences of a large move towards central clearing. Around a quarter of the OTC derivatives market (by notional) is already centrally cleared and this fraction is likely to increase significantly over the coming years. Clearly. a close look at the strengths and weaknesses of central clearing is relevant.

Systemic Risk

One of the key concerns over the global OTC derivatives market has always been systemic risk. Systemic risk does not have a firm definition but is essentially financial system instability exacerbated by distress of financial intermediaries. In the context of counterparty risk, systemic risk could arise from the failure of a large financial institution or intermediary and the inevitable knock-on effects. creating financial problems for other market participants, These institutions may then, in turn, fail and continue the "domino effect".

Systemic risk will therefore generally involve some initial spark followed by a proceeding chain reaction, potentially leading to some sort of explosion in financial markets, Thus, in order to control systemic risk, one can either minimise the chance of the initial spark, attempt to ensure that the chain reaction does not occur or simply plan that the explosion is controlled and the resulting damage limited.

Reducing the default risk of large. important market participants reduces the possibility of an initial spark caused by one of them failing. Capital regulation and prudential supervision can contribute to this but there is a balance between reduction of default risk and encouraging financial firms (and the economy) to grow and prosper. DPCs and monolines, discussed previously, are good examples of this balance. Placing very stringent capital and operational limitations on such an entity will make it extremely creditworthy and yet simultaneously make it impossible to generate the returns required to function profitably as a corporation. However, without the correct management and regulation, ultimately, even financial institutions that once seemed like fortresses can collapse.

Given that firms will inevitably fail, having efficient market mechanisms and structures for containing the failure of key firms in place and absorbing a large shock is key. OTC derivatives markets have netting, collateral and other methods to minimise counterparty risk. However, such aspects create more complexity and may catalyse growth to a level that would never have otherwise been possible. Hence, it can be argued that initiatives to stifle a chain reaction may achieve precisely the opposite and create the catalyst (such as many large exposures supported by a complex web of collateral) to cause the explosion.

The ultimate solution to systemic risk may therefore be simply to have the means in place to manage periodic explosions in a controlled manner, which is the role of a CCP. If there is a failure of key market participant then the CCP will guarantee all the contracts of that counterparty executed through them as "clearing members." This will mitigate concerns faced by institutions and prevent any extreme actions by those institutions that could worsen the crisis. Any unexpected losses[1] caused by the failure of one or more counterparties will be shared amongst all members of the CCP (just as insurance losses are essentially shared by all policyholders) rather than being concentrated within a smaller number of institutions that may be heavily exposed to the failing counterparty. This "loss mutualisation" is a key component as it mitigates systemic risk and prevents a domino effect.

[1] Meaning those above a certain level that will be discussed later.

The Impact of the Crisis

In 2007, a US housing crisis led to a credit crisis, which caused the failures of large financial institutions and a severe economic downturn. Authorities had to make key decisions over failing institutions such as Lehman and AIG with a very opaque view of the situations the firms were in and the potential knock-on impact of any decisions made. The dramatic increase in counterparty risk in the early stages of the global financial crisis and the realisation that no counterparty was immune to potential financial distress brought many calls for a dramatic solution. Regulators were swift to act and CCPs seemed to emerge as a panacea for solving counterparty risk problems.

For example, US policymakers fast-tracked a number of changes to improve the derivatives markets, in particular with respect to CDS contracts. In May 2009, the Obama administration (through the US treasury) proposed a new framework for greater market regulation and oversight of the OTC derivatives market.[2] This framework, broadly speaking, mandated central clearing of all CDS transactions as well as prudent regulation of CDS market participants and increased transparency. At the end of 2008, the SEC approved a series of temporary conditional exceptions that allowed trading index CDSs through CCPs without the delays and hurdles that full regulation would create. The SEC did this in the belief that a CCP can reduce systemic risk, operational risks, market manipulation and fraud and contribute to overall market stability.[3]

By late 2010, both EU (via the European Commission) and US regulation (via the Dodd–Frank Wall Street Reform and Consumer Protection Act) had published formal legislative proposals that all standardised OTC derivatives should be cleared through CCPs. Basel III capital requirements, first published in 2009, incentives central clearing through relatively low capital requirements for CCP-cleared (as opposed to bilateral) trades. Whilst there will be technical standards and implementation details to be decided, it appears clear that the OTC derivatives market will move significantly more towards central clearing in the coming years.

[2] Press Release, US Department of Treasury, *Regulatory Reform OTC Derivatives*, 13th May 2009 (http://ustreas.gov/press/releases/tg129.htm).

[3] See http://sec.gov/rules/exorders/2008/34-59164.pdf

CCPs in Perspective

Most CCPs were originally created by the members of futures exchanges to manage default risk more efficiently and were not designed as the saviour of the global financial system. However, they operated well in the otherwise chaotic financial markets around the Lehman default and this appears to be a key driver for the logic behind regulators and policymakers deciding that the future OTC derivatives market needs to be mainly CCP-based. However, many market participants do not believe that CCPs will definitely make financial markets a better and safer place in the long-term.

The aim of this chapter is to present a critical analysis of CCPs, without attempting the futile task of predicting what their overall impact on future financial markets will be. Some important points to explore are:

- A CCP cannot make counterparty risk disappear. It can only centralise it in a single place and convert it into different forms of risk such as operational and liquidity.
- As with most risk mitigation, for every advantage of a CCP, there are related disadvantages. The benefit of a CCP is not to be a panacea but rather to lead to more advantages than disadvantages.
- CCPs can *reduce* systemic risk (which is well documented) but can also *increase* it.
- There are likely to be unintended consequences of the expanded use of CCPs, which are hard to predict *a priori*.

CCPs must have a fine-tuned structure with respect to collateralisation, settlement and risk management and obviously must be extremely unlikely to fail. Furthermore, the very nature of what a CCP does creates potentially dangerous problems such as moral hazard and adverse selection. Finally, the failure of a CCP will clearly be an extremely problematic event, potentially far worse than the failure of any other financial institution.

A large part of the motivation for the regulatory push towards central clearing arises from strong concerns regarding the regulation and practices within the CDS market. However, such concerns must not be overstated due to the unprecedented nature of the crisis. The start of the crisis was largely the result of systematic mispricing of mortgage-related debt and not directly due to the growth of the credit derivatives market. The systemic failure of counterparty risk in CDSs occurred only because of regulated financial guarantors (such as AIG) selling risky protection on assets such as MBS. AIG's excessive risk-taking via CDSs was part of a broader problem related to seeking returns from mispriced mortgages and mortgage-backed securities.

Function of a CCP

As mentioned already, a CCP is traditionally an entity set up to manage the counterparty risk that exists on an exchange. All derivatives exchanges have adopted some form of CCP. Clearing is what takes place between trade execution and trade settlement (when all legal obligations have been made). When trading a derivative, the counterparties agree to fulfil specific obligations to each other. By interposing itself between two counterparties, which are clearing members, a CCP assumes all such contractual rights and responsibilities. As a result, an institution no longer needs to be concerned about the credit quality of its counterparty, indeed the counterparty to all intents and purposes is the CCP.

The legal process whereby the CCP is positioned between buyer and seller is known as novation. Novation is the replacement of one contract with one or more other contracts. This will not happen immediately and therefore bilateral counterparty risk will exist for a short period.[4] The viability of novation depends on the legal enforceability of the new contracts and the certainty that the original parties are not legally obligated to each other once the novation is complete. Because of novation, the contract between the original parties ceases to exist and they therefore do not have counterparty risk to one another. Because it stands between market buyers and sellers, the CCP bears no net market risk, which remains with the original party to each trade. On the other hand, it does take the counterparty risk, which is centralised in the CCP structure. A CCP will attempt to mitigate most of this risk by demanding financial resources from its members that cover the potential losses in the event they default.

CCPs use collateral to mitigate the counterparty risk they face although this is normally referred to as "margin". In addition to the standard "variation margin" that covers the change in the valuation of the relevant positions as with the collateral in a standard CSA,[5] CCPs will require "initial margin" that overcollateralises[6] the counterparty risk they face. Initial margin exists for the life of the trade and can be increased or reduced depending on market conditions and the remaining risk. Members will also need to commit other financial resources to the CCP such as through what is often known as the '"default fund" or '"reserve fund". The initial margin and reserve fund contributions required globally for the future move of OTC derivatives to CCPs are significant and comparable with the increases in capital required due to regulatory changes since the crisis. Clearing members may also partially own the CCP.

By taking initial margin and overcollaterlising positions, a CCP transforms some counterparty risk into other forms of risk. Most obviously, liquidity risk arises due to the nature of margining but operational and legal risks also exist.

The ideal way for CCP members to contribute financial resources is a "defaulter pays" approach. This would mean that a member would contribute all the necessary funds to pay for their own potential future default. This is impractical though because it would require costly initial margin (and reserve fund contributions) to be set at an impractically high level. For this reason, the purpose of initial margin is to cover most (around 99% or more as discussed later) scenarios where a member would default. This leaves a small chance of losses not following the "defaulter pays" approach being home by the other clearing members. Whilst this is a consequence of the need to reduce systemic risk, it creates problems such as moral hazard, which must be considered carefully.

Multilateral Netting

The primary advantage of a CCP is multilateral netting, which can alleviate systemic risk by reducing exposures more than in bilateral markets. Let us compare the different netting schemes in relation to an example set of exposures between three counterparties as shown in Figure 12-1:

- *No netting.* Default of any institution will give rise to losses of 3 and 5 for the remaining institutions. For example, a default of A will cause a loss of 5 for B and

[4] Currently a few days but in future likely to be just a few hours as most CCPs aim to clear trades by the end of the relevant business day.

[5] This acts like a two-way CSA with a zero threshold for both counterparties.

[6] Initial margin therefore is analogous to an independent amount.

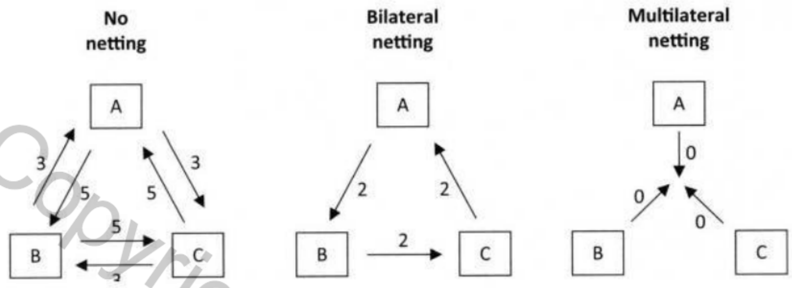


FIGURE 12-1 Comparison of netting schemes. An arrow indicates the direction of money owed (exposure) so that—under no netting, for example—entity A has an exposure of 3 to entity B whilst B has an exposure of 5 to A.

3 for C whilst A will still claim a total amount of 8 owed to them.

- *Bilateral netting.* Default of A will cause a reduced loss of only 2 for B whilst C will suffer no loss at all (since they owe A).
- *Multilateral netting*. No institution is exposed to another since none has any outstanding exposure. Nor does the CCP have risk (in this stylised example) since all positions net.[7]

Whilst multilateral netting is clearly more beneficial when all trades are covered, in reality fragmentation will be a problem. In such cases, one has to assess the benefits of multilaterally netting a subset of trades against losing the bilateral benefits that these trades have. A simple and intuitive quantitative treatment of the benefits of a CCP is given by Duffie and Zhu (2009). Their results are based on considering the netting benefit (based on the measure known as EPE for trading a single class of contracts through a CCP as opposed to bilateral clearing. They show. using a simple model,[8] the required number of dealers trading through the CCP for a single asset class to achieve netting reduction. The results are plotted in Figure 12-2 as a function of correlation and number of asset classes. For example, for four uncorrelated asset classes, there must be at least 15 dealers to make clearing a single asset class through the CCP valid.[9]

The above example assumes equal distribution of exposure across asset classes. Duffie and Zhu also consider a non-homogeneous case and derive an expression[10] for the fraction of a dealer's exposure that must be concentrated in a particular asset to make a CCP for that asset class viable. This fraction is shown in Figure 12-3. For example, with 10 dealers, using a CCP for a given class of derivatives will be effective only if three-quarters of the dealers' bilaterally netted exposure resides in that class of products.

The Duffie and Zhu results illustrate that achieving overall netting benefits from central clearing (compared to bilateral trading) is not a foregone conclusion. Increased netting benefits can only be achieved by a relatively small[11] number of CCPs clearing a relatively large volume of transactions. However, it could be argued that the success of central clearing does not depend on the reduction of exposure but rather the control of systemic risk (although the two are partly related).

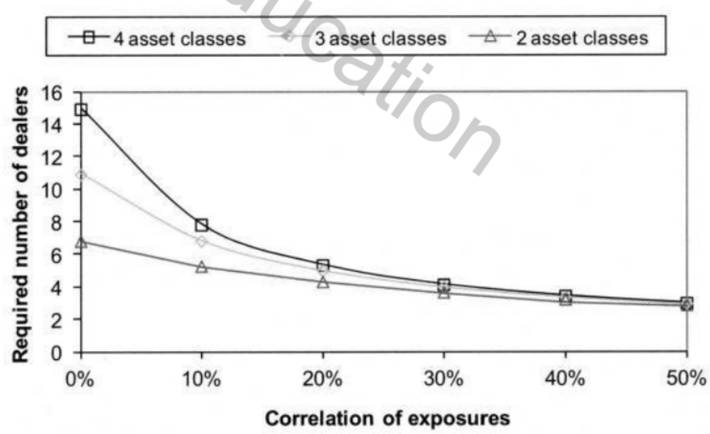


FIGURE 12-2 Required number of dealers for a single asset class CCP to improve netting efficiency calculated using the formula of Duffie and Zhu (2009).

[7] In the example, the matching exposures may be by chance or may arise due to perfectly matching (mirror) trades. In the former case, the CCP will have risk but the point is to illustrate that multilateral netting decreases exposure still further compared with bilateral netting.

[8] Simplifying assumptions of symmetry and equal variance of exposure are used in this case.

[9] Interestingly, the impact of correlation between asset classes makes a CCP more effective since bilateral netting is less effective in this case.

[10] This assumes independence between asset classes.

[11] See also the later comments on interoperability.

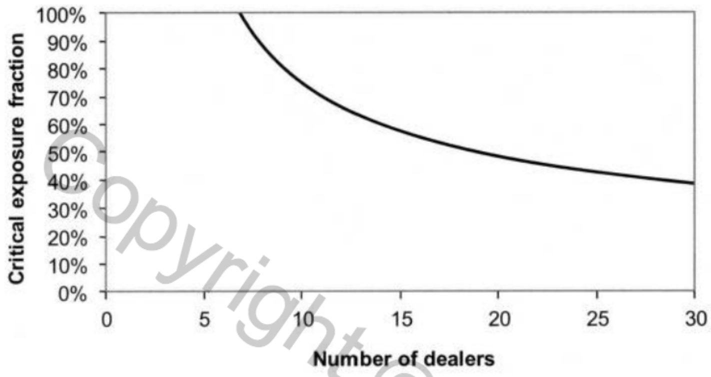


FIGURE 12-3 Required fraction of exposure attributed to a single asset class ("critical exposure fraction") to make a CCP for that asset class effective. The results as a function of the number of dealers are calculated using the formula of Duffie and Zhu (2009).

Given CCPs are unlikely to reduce exposure, the collateral requirements for central clearing are likely to be high. For example, Singh (2010) estimates $2 trillion globally. However, it is argued (e.g., Milne, 2012) that the overall private cost of collateral is low due to Modigliani-Miller-type arguments (firms holding extra collateral can simply borrow more) and the social cost is lower still.

How Many CCPs?

A further result of the Duffie and Zhu study mentioned above is that, not surprisingly, if a CCP is viable then it is inefficient to have more than one CCP in the market. In reality, many competing forces will define the number of CCPs that exist globally. Factors favouring a small number of CCPs are the multilateral netting benefits and other economies of scale. There is likely to be a degree of consolidation of CCPs to reflect these forces. Often a CCP may start with a single goal and therefore be focused in terms of region and asset classes. Growth will naturally involve expanding the geographical base, markets and products covered. The move to central clearing is naturally limited in the early stages due to the lack of benefit until a critical mass of products is cleared.

However, there are a number of factors that will push the total number of CCPs upwards. Jurisdictional fragmentation is the first obvious hurdle to CCP consolidation. Regulators in several major jurisdictions have made it clear that products traded in them or by firms located there must be cleared there. CCPs clearing two or more different types of products are also problematic. The less liquid product takes longer to closeout. This could create a priority where members predominantly trading in the more liquid products (for example, interest rate swaps compared to credit default swaps) would benefit by having an effective first claim on the initial margin and reserve fund of the defaulted member. This is one reason why some CCPs have typically been focused around a single asset class and even product type.

The number of CCPs also depends on the need for competition between CCPs. In one sense, competition may be preferable, as market forces will determine the CCP landscape and costs (via margins) may be competitive. However, too much competition may be counterproductive for the overall risks that CCPs represent. The potential danger of competition for risk assessment was provided by ratings agencies. Up to 2007, a plethora of ever-more-complex products were given good-quality ratings driven by the fact that ratings agencies were essentially paid for giving such ratings and competing with one another. Competition between CCPs can be dangerous and perhaps a preferable solution would be a single monopolistic CCP that concentrates on strong risk management and not on providing attractive costs to potential members. Indeed, this model has existed in Brazil for a number of years.[12]

The balance of forces described above means that it seems likely that some sort of equilibrium may be established with the total number of CCPs globally being in double figures but potentially not too far into double figures. It seems reasonable that the financial markets would be best served via a moderate number of CCPs, which are large enough to offer good product coverage but not so large that they create monopolistic issues, severe systemic risk or geopolitical problems.

Coverage of CCPs

CCPs have historically been closely affiliated with exchanges and have therefore been responsible for the clearing of standard exchange-traded products. Only a few CCPs have offered central clearing services for OTC derivatives (e.g., LCH Clearnet's SwapClear). Central

[12] In Brazil, only a single clearing house exists (BM&F Bovespa) which clears some OTC derivatives and also acts as a central securities repository.

clearing or OTC products presents a number of problems due to a number of aspects, and many products are likely to remain OTC for the foreseeable future. Regarding the types of product tradable through a CCP, the considerations are:

- *Standardisation*. Given the nature of clearing a trade, which involves responsibility over contractual payments and valuation for the purposes of margin calculation, products must be relatively standard. For example, in 2009 there was a standardisation of CDS contracts, which was a prerequisite to any migration to CCPs.
- *Complexity*. Complex derivatives are more problematic for CCPs since their contractual features and valuation will be less straightforward. Note that a complex or exotic derivative may be standardised but still problematic for these reasons.
- *Liquidity*. More illiquid products have less accurate pricing information and typically less historical data for calibrating risk models (for the purposes of initial margin calculation). It is also more costly to replace illiquid trades in the event a CCP member defaults. Note that the liquidity of a trade may well be transient, i.e., it may be liquid at inception but illiquid later. This could be due to either contractual terms or market factors.
- *Wrong-way risk*. Wrong-way risk products are more dangerous for CCPs since there is expected to be a larger exposure in the event of default of the clearing member. This is the main reason why CDS products are more complex to clear than interest rate swap products (even if the additional volatility in the CDS spreads has been accounted for).

The result of the above is that currently only a relatively small number of OTC derivatives products are suitable for central clearing. However, given the size of the interest rate swap market, just clearing this product type will have a very significant impact. CDS indices are also currently cleared and, whilst they are standardised, they are more complex, illiquid and have wrong-way risk. Single-name CDSs will be even more problematic due to their inherent "jump to default" risk. Finally, the OTC derivatives that contributed so significantly to the global financial crisis, namely tranches of bespoke credit portfolios, are extremely difficult to imagine within a CCP environment.

With the above comments in mind, there are two schools of thought with respect to centrally cleared products. The first is that by clearing products that are relatively easy to deal with, such as interest rate swaps and the simpler (index) credit default swaps, a large fraction of OTC derivatives notional can be cleared with the residual remaining forever OTC. An alternative view could be that the products that are difficult to clear are precisely the biggest danger in terms of potential future crises. This would imply that a long-term aim should be to clear as many OTC derivatives as possible. However, it is hard to envisage exotic, illiquid or highly structured OTC derivatives being centrally cleared in the near future, or ever.

An example of the opposing views was illustrated with respect to FX products. Market participants have resisted the idea of central clearing for FX products on the basis that the exposures are generally much smaller and settlement risk is more problematic. However, Duffie (2011) argues against this, pointing out that whilst many contracts are short dated, some FX trades have significant credit exposure because FX rates can be volatile, have fat tails and be linked strongly to sovereign risk.

We noted above that a contract could be standardised but still complex. This is a particularly important point to note for credit derivatives products. A single-name CDS is standard and has a payoff that appears quite simple: it is a swap where one party pays a premium in return for receiving a contingent payment if a pre-specified credit event occurs. However, we could also describe this product as being an American-style out-of-the-money barrier option.[13] A CDS index would then be a portfolio of out-of-the-money barrier options. Furthermore, CDS contracts are relatively illiquid (compared to IRS, for example) and become even more so in a crisis. The standardisation of a product does not make it less complex or more liquid.

Pirrong (2009) argues that asymmetric information costs will be higher in centrally cleared markets compared to bilateral ones, especially for exotic products traded by complex, opaque intermediaries. This is argued to be due to the specialisation of dealers with respect to valuing exotic derivatives together and the fact that dealers are more effective at, and have more incentive for, good monitoring and pricing of counterparty risk compared to a CCP. Market participants trading with a CCP may be

[13] Out-of-the-money because the credit event is usually relatively unlikely; barrier because default is similar to the crossing of a barrier, and American-style because default can occur at any time.

incentivised to create larger positions than they would otherwise like to, or even be able to, occupy. Put another way, a CCP may therefore suffer from a form of "winner's curse". Such a phenomenon is well known in insurance markets where an insurance company will naturally end up with more risk due to policyholders automatically finding the cheapest premiums[14] given their circumstances.

LOGISTICS OF CENTRAL CLEARING

Clearing Members

From the point of view of trading through a CCP, one can consider three types of participant:

- General clearing member (GCM)—a member of the CCP who is able to clear third parties as well as their own trades.
- Individual clearing member (ICM)—a member of the CCP who clears only their own trades.
- Non-clearing member (NCM)—an institution having no relationship with the CCP but which can trade through a GCM.

These relationships are illustrated in Figure 12-4.

A GCM or ICM will typically be a large bank or dealer who has a large number of counterparties. An NCM is more obviously characterised by an end-user of OTC derivatives that may channel most or all of its trades through a single counterparty. By this counterparty being a GCM,

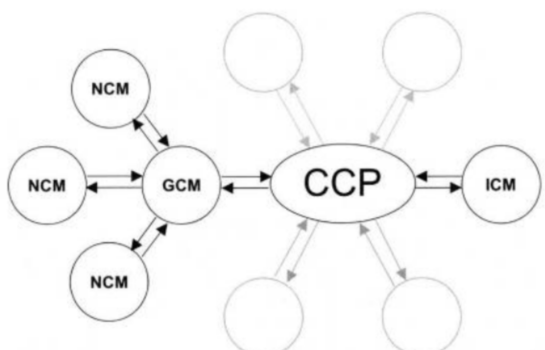


FIGURE 12-4 Illustration of the relationship between CCPs and clearing members.

[14] For example, an insurance company specialising in insuring drivers with motoring convictions may be able to charge higher premiums but market forces will ensure that they end up with more risky drivers on their books.

the end-user can gain benefits from central clearing even though they are not a clearing member. Whilst these are the more obvious roles of members, they are not the only characterisations: for example, a GCM may in fact be another CCP and an NCM may be a smaller bank.

Variation Margin

Variation margin is a simple concept, which is an adjustment for the change in valuation of the relevant positions at periodic intervals of at least daily frequency. A CCP may also make intraday margin calls if large price movements threaten to exhaust margin funds in a clearing member's account. Such practices are becoming increasingly common and are supported by technology advances. Valuation is typically straightforward since a prerequisite for clearing is that the underlying trades are standardised. Hence, there is little subjectivity over variation margin amounts.[15] Variation margin will typical be cash in major currencies (or alternatively highly liquid securities). As a counterparty to all trades, CCPs are calculation agents, valuing all positions and collecting or paying respective margins.

Impact of Default of a CCP Member

In the event of a default, the responsibilities of a CCP are broadly:

- Auctioning the positions of the defaulted member.
- Transferring client positions to surviving clearing members ("porting", see later).
- Allocation of any excess losses to surviving clearing members.

The viability of a CCP depends on its ability to withstand the default of one or more clearing members. As in bilateral markets, the first line of defence for a CCP in such a scenario will be to essentially closeout all positions for the member in question and apply netting of such positions. However, as mentioned previously, we note that CCPs will probably auction positions amongst surviving members rather than terminate them completely.[16] In an ideal world,

[15] Although aspects such as OIS discounting and funding have made even the valuation of standard interest rate swaps a complex problem.

[16] In the case positions were terminated, members would be likely to attempt to execute replacement trades, which then may be subject to central clearing by the CCP. The auction is a quicker and more efficient way to achieve this.

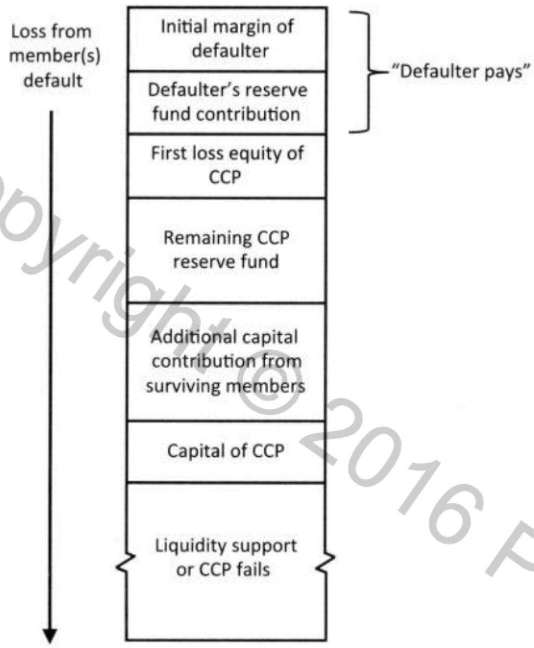


FIGURE 12-5 Illustration of a typical loss waterfall defining the way in which the default of one or more CCP members is absorbed.

the termination of positions with a defaulting member would be matched perfectly by the variation margin with respect to that member. In reality, whilst the variation margin should cover any associated losses at a time just prior to the default event, the actual period of time the CCP is exposed to the positions will be longer than this. Hence, there is a need for additional financial resources to absorb further losses. A CPP will have several layers in order to absorb such losses and a typical "loss waterfall" is represented in Figure 12-5.

In most cases, the initial margin is intended to cover the losses from a defaulting member. If this is not sufficient then losses are absorbed from the defaulter's contribution to the reserve fund. Within these amounts, the "defaulter pays" approach, where the defaulting member covers all losses, is being followed.

In extreme scenarios, where the initial margin and reserve fund contributions of the defaulted member(s) have been exhausted, further losses may be taken from the first loss equity of the CCP (for example, current annual profits). After this point, the financial resources contributed by other members via the remaining reserve fund contributions are used. This is the point at which moral hazard appears as other members are paying for the failure of the defaulted member. Indeed, it is possible for a clearing member to lose some or their reserve fund when another member, with whom they have avoided trading, defaults.

Losses wiping out the entire reserve fund of a CCP are clearly required to be exceptionally unlikely.[17] However, if this does happen then the surviving members of the CCP are required to commit some additional "capital" to support the CCP. This contribution is not unlimited and is capped (often in relation to a member's initial reserve fund contribution) as a means to mitigate moral hazard. After the cap is reached then the remaining capital of the CCP would be used to cover losses.

At this point, assuming losses still persist, the CCP will fail unless they receive some external liquidity support (via a bailout from a central bank, for example). We should note that, in order to reach the bottom of the loss waterfall, many layers of financial support must be eroded. Hence, although unquantifiable to any relative precision, this will be an extremely low probability event.[18]

We will now discuss the key elements of the loss waterfall, namely the initial margin and reserve fund in more detail.

Initial Margin

Initial margin is, by contrast to variation margin, much more complex and its magnitude is highly subjective. Initial margin is a buffer posted to cover the closing out of positions without loss to the CCP in a worst-case scenario. Initial margin requirements need to be set carefully depending on the trade in question, any existing trades and the required confidence level of the CCP. A final important consideration is the assumed time-period that the CCP would be exposed to the value of a member's positions after their default event (including any grace period). A confidence level of 99% and time-period of around 5 days are typical.

Initial margin depends primarily on the market risk of the centrally cleared trades and only a small component. if any, is linked to the credit quality of the clearing member. Since the posting of cash or high-quality assets imposes higher costs for members, margin requirements must not be too conservative as this may reduce trading volumes

[17] Indeed reserve funds are typically calibrated based on the assumed default of the two largest members.

[18] However, never as low as we think.

through the CCP. On the other hand, they need to be large enough to provide the right level of risk mitigation. A CCP must balance the need to be competitive by incentivising central clearing (low margins) with maximising their own creditworthiness (high margins). i.e., to what extent the default of a member can be absorbed.

The following components are important considerations in deciding on initial margin:

- *Volatility.* The most obvious aspect is the volatility of the trade in question. This is driven by the volatility of the underlying market variable(s) and the maturity of the trade.
- *Tail risk*. Whilst volatility typically measures continuous price variability, some products such as CDSs can suffer from tail risk due to jumps or gaps in the underlying market variables.
- *Dependency*. Since CCP members typically hold a variety of trades with a CCP, it is important to understand the offsetting nature of such trades. If correlation between the price moves of different trades is small then clearly the overall portfolio is less risky and the benefit of this is that less margin needs to be charged. Gemmill (1994) has illustrated the diversification offered to a CCP from clearing several markets that are not highly correlated. Indeed, CCP members will be actively looking to receive the benefits of such favourable dependencies in the form of lower initial margins. CCPS tend to offer benefit from same-asset-class trades but less so from cross-asset-class interactions, although this is likely to become more common.

It should be noted that the above calculations, ideally, should all be made under the assumption that the member in question is in default. For example, if a member's default is expected to produce particular volatility in the CDS market then this volatility is the appropriate measure to use when setting initial margin.[19] This is essentially accounting for wrong-way risk.

Historically, systems such as SPAN[20] have been used for initial margin, evolving risk factors combinatorially based on movements in either direction based on a certain confidence level (this is similar to some methods now used for stress-testing purposes, for example). The worse-case scenario is normally used to define the initial margin. Whilst such methods work well and are tractable for simple portfolios such as futures and options, they have severe drawbacks. Mostly notably, they do not scale well to a large number of dimensions (as the number of combinations of moves grows exponentially as 2^n where n is the number of dimensions). They also do not clearly attach an underlying probability to the scenario defining the initial margin.

More advanced initial margin calculations follow value-at-risk approaches. One lesson from many years of application of VAR methodologies is that "reasonable" losses can be quantified with some success, whereas more severe losses fall outside the abilities of quantitative approaches. Not surprisingly, this can lead to an underestimate of required margin. For example, Figlewski (1984) has shown for equity markets that for a confidence level of 95% such an approach can work reasonably well whereas for a higher confidence level of 99% the empirical margin requirements are much higher than those predicted using normal assumptions.[21] Bates and Craine (1999) showed that following the 1987 crash, the expected losses conditional on a margin call being breached increased by an order of magnitude. Clearly, it is important to attempt to incorporate any "fat tail" effects in the calculation of margins. Beyond the univariate assumptions applied to margin calculations, it is important to consider the multivariate behaviour of different positions (for example, in different currencies) to understand how the total margin would be reduced due to netting effects.

VAR approaches for initial margin will aim to calculate the worst-case loss for a set of trades of any dimensionality and accounting for all cross-dependencies. This will typically rely on a certain window of historical data for the calibration of parameters such as volatility and correlation. Such an approach is scalable to any number of products and risk factors and can allow initial margins to be calculated for large multi-asset portfolios, giving all the relevant netting benefits. Indeed, CCPS are looking in some cases

[19] Alternatively, this could be a component of the reserve fund contribution although then the confidence level for initial margin will be implicitly lowered.

[20] Standard Portfolio Analysis of Risk, used historically, for example, by US futures clearing houses and the London Clearing House.

[21] A more complex approach such as extreme value theory (EVT) potentially provides a more sophisticated way to attempt to capture the possibility of extreme price movements. A weakness of EVT is that it requires significant historical data whereas most CCPs prefer to set margin based on a small and recent data sample. EVT may capture potential extreme moves and will therefore lead to higher initial margins.

to offer such netting benefits by combining margins for different product types.[22] However, whilst multidimensional modelling of risk factors can lead to increased benefits from margin offsets, it does increase the model risk. Dependencies of financial variables are notoriously hard to model, especially in certain cases such as involving credit spreads.

Finally, we note that on top of the univariate and multivariate challenges in computing initial margin, there is a procyclicality problem. In benign markets, volatility and correlations will be low and calculated margins therefore smaller. However, this means that, in more volatile times, margins will be insufficient to provide the coverage that the confidence level and time period intended. Since volatility can increase and dependencies change rapidly, it may be necessary to increase margins in periods that are more turbulent. Indeed, the fact that this practice has become *de rigueur* has made the term "initial margin" somewhat confusing since it implies that this margin will be fixed over the lifetime of a transaction where, in reality, market conditions may cause it to increase. The above quantitative difficulties can be resolved by being conservative, especially during periods of low market volatility. Such an approach would mirror the stressed VAR and EPE requirements for capital under Basel III, where the model must always be calibrated to a dataset including period of stress.

However, initial margin should be at a level to cover all but extreme price movements, but not so high as to damage market liquidity and/or discourage the use of the CCP. High margins have been shown empirically to have a detrimental impact on trading volumes (e.g., see Harlzmark, 1986; Hardouvelis and Kim, 1995).

Whilst it may seem reasonable for a CCP to call for more initial margin during turbulent periods, by doing this the CCP creates systemic risk by, to put it crudely, sucking liquidity out of the market at a critical time. Brady (1988) discusses the crash of 1987 and its impact on some clearing houses arising in an extreme market event with associated liquidity problems. Whilst a moderate increase in margins during more volatile periods may not be unreasonable, a significant hike in margins close to a crisis could have a severely destabilising impact on prices and create unpleasant knock-on effects. Whilst CCPs are generally intended to reduce systemic risk, increasing initial margins is one way in which they may achieve the reverse. The analogy of using stressed market data as a basis for VAR calculations would be that CCP in initial margins should upfront contain a contribution relating to a more volatile period to reduce the need to call for additional margin later.

Initial margin will normally be required in cash or in some cases other very liquid securities. Non-cash margins posted with CCPs will typically not be rehypothecated, as this creates the need for more liquid assets and potentially puts a strain on liquidity.

A critical point to emphasise is that CCPs do not vary initial margin significantly based on the credit quality of the clearing member.[23] Therefore, members of a CCP are essentially treated equally. This has an obvious implication that CCP members must be rather similar in credit quality. Even then, there are asymmetric information problems since weaker members will be gaining at the expense of stronger members. Indeed, Pirrong (2000) argues that the delay in adopting central clearing on certain exchanges was related to stronger credit quality members not wishing to subsidise weaker ones.

Reserve Funds, Capital Calls and Loss Mutualisation

Whilst keeping the likelihood of exceeding margin over a single day to a high confidence level (such as 99%) is viable, breaches will always be possible. CCPs don't only focus on having margin requirements that cover losses in all but the most extreme cases. They aim to ensure that there is adequate coverage of losses due to the default of a member following a margin-depleting price move. The ability of a CCP to survive such extreme losses, potentially arising from default or several members, is critical.

Another important aspect of CCPs is the reserve fund that has been accumulated over time by initial and ongoing contributions from clearing members and/or built up from other sources such as CCP profits. A determination of the correct reserve fund is significantly more difficult than the

[22] See, for example, "Clearinghouses seek to merge margin accounts for CDS clients", *Dow Jones Newswires,* 7th October 2011.

[23] Some CCPs do base margins partially on credit ratings, for example by requiring more when a member's rating falls below a certain level. However, this is clearly problematic since credit ratings are imprecise and granular measures of credit quality and such triggers requiring more margins are well known to be potentially destabilising and create systemic risk.

already complex initial margin calculation outlined above. By their very nature, losses hitting the reserve fund are infrequent and typically represent a 1% or lower probability **and** the default of at least one clearing member. Calculating the conditional exposures in such scenarios is plagued by problems such as fat tail events, complex dependencies and wrong-way risk. For these reasons, CCPs typically calibrate the size of a reserve fund more qualitatively via stress tests typically framed in terms of the number of defaults a CCP can withstand (two large defaults, for example). However, the true probability of a CCP exhausting their reserve fund is very difficult to quantify with any accuracy as it is linked to events involving default of more than one clearing member together with extreme moves in the positions of such members.

Losses above the defaulting member's reserve fund contribution will begin to hit other clearing members. This loss mutualisation is a key point since it spreads extreme losses from the failure of a single counterparty across all other clearing members. This has the potential to ameliorate any systemic problems arising in bilateral markets due to an institution heavily exposed to a defaulted counterparty. Loss mutualisation (along with margining requirements) completes the process of homogenisation of counterparty risk across all clearing members.

When the reserve fund is exhausted, there will be additional contributions or "capital calls" from remaining clearing members. Such contributions are capped, which reduces moral hazard problems. However, this means that a CCP could become insolvent and without some external injection of capital would fail.

Interoperability

As mentioned above, it seems likely that jurisdictional issues and potential product segregation will create a relatively large number of CCPs. Clearing the same product in multiple CCPs causes fragmentation and limits the benefits of clearing, especially in terms of multilateral netting. Local regulators requiring clearing by CCPs in their own region will also create a need for interoperability. An obvious way to improve such a situation is to link CCPs. Such interoperability would have the advantage of increasing multilateral netting benefits and therefore giving clearing members lower initial margin requirements thanks to the offsetting nature of positions even though they trade those positions through more than one CCP.

Interoperability will firstly require cooperation between CCPs on the initial margin required, taking into account the netting benefits from the linkage of the CCPs. For example, a bank may clear CDS through CCPI and IRS through CCP2. The CCPs must decide on the total initial margin required for these two sets of trades, knowing that they are not perfectly correlated. Then there is the question of where to hold this margin and how to allocate it in the event the bank were to fail. Even theoretically, such loss sharing is not an obvious calculation. Linking CCPs across different jurisdictions and regulatory regimes will also clearly be difficult and require some harmonisation of regulatory rules and cross-border bankruptcy rules.

Interoperability can mitigate problems due to multiple CCPs and increase the benefits of clearing but would expose CCPs to each other's failure and create a more interconnected CCP landscape.

Non-Clearing Members and End-Users

One of the key challenges of moving OTC derivatives to CCPs is persuading "end-users" such as hedge funds, asset managers, insurance companies, corporates and sovereigns to move their positions with dealers. Firstly, end-users often have limited margin (collateral) arrangements with banks and may benefit from not posting margin or posting only limited margin (due to a threshold in the CSA). They may also be able to post illiquid margin (for example, corporate bonds, gold and in extreme circumstances aeroplanes). In order to trade through a CCP, an NCM will either have to commit to frequent posting of liquid margin as required by the CCP or will have to rely on their GCM to provide some sort of "collateral upgrade" service with respect to margin transformation. For example, it is possible that an NCM may post less liquid margin to a GCM and then the NCM will in turn post the required more liquid margin to the CCP. Already the business of collateral conversions has begun in earnest with banks offering the service for their clearing clients via their securities lending and repo desks. Obviously, the GCM would charge for such a service, probably via demanding a significant haircut on the more illiquid margin.

The extreme case of the above is margin lending, where a GCM or other third party essentially funds the NCM for the margin that they need to post. Indeed, Albanese *et al.* (2011) have proposed this as a solution to the significant initial margin funding requirements for central clearing.

However, we note that such a liquidity provision creates potentially dangerous problems because the fee for margin lending will increase as the credit quality of the borrower increases. Furthermore, the failure of entities such as conduits and SIV (structured investment vehicles)[24] during the crisis suggests that, in extreme markets, margin-lending mechanisms could simply freeze completely.

There are arguments that some end-users should be excused from CCP clearing on the basis that it will be too difficult and expensive to hedge their genuine risk profiles. For example, US regulation provides exemptions for hedgers who are not dealers or "major swap market participants." At the time of writing, it is not completely clear what organisations are covered by such a definition.

Another important consideration is issues that arise over segregation of margin. When customer collateral is commingled in "omnibus" accounts, it is at risk in the event of the clearing member's default. Furthermore, clearing members may be able to utilise margins of non-defaulting customers to offset the obligations of another defaulting customer. This latter effect can be seen as another form of moral hazard where good credit quality clients of a clearing member are in danger of experiencing losses due to the failure of weaker clients. Segregation of margin reduces the risk that a client will lose some or their entire margin in the event of a default (of their clearing member or of another client of their clearing member). There is no free lunch, however, as greater segregation is more costly, both operationally and in terms of the requirement that reserve funds be larger (to cover the losses that would otherwise be taken from surviving institutions margin). There is also the question of the legal enforcement of segregation across jurisdictions.

The final issue is that, in the event of default of their clearing member, clients not only want not to lose their margin. but they also require the ability to "port" their positions to a surviving clearing member (Figure 12-6). This is clearly highly preferable to having all positions closed out or distributed via multiple clearing members through an auction (which would lead to the requirement to post more margin due to the loss of netting benefits). Whilst having all positions moved to a single clearing member is useful, there is the question of whether, in the potentially turbulent markets around a clearing member default, another member will require significant additional margin or simply refuse to accept some or all of the trades.

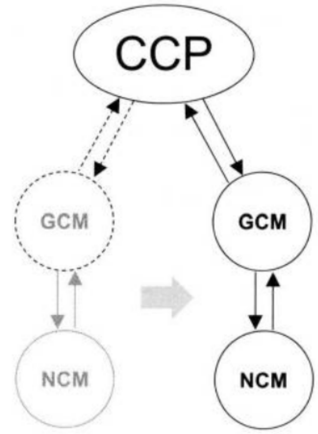


FIGURE 12-6 Illustration of the "porting" of non-clearing member positions from a defaulting general clearing member to a surviving general clearing member.

Note that greater segregation and the ability to port positions creates more moral hazard since a client is less likely to monitor the health of their clearing member.

ANALYSIS OF THE IMPACT AND BENEFITS OF CCPs

In the description of the mechanics of CCP trading, advantages and disadvantages emerge. We have also described the creation of effects such as moral hazard and adverse selection. These are not insurmountable disadvantages (after all, the insurance market operates despite such problems), but may lead to unintended consequences. We now look at the strengths and weaknesses of CCPs in more detail.

The Advantage of Centralised Clearing

The advantages of trading through a CCP are:

- *Multilateral netting*. Contracts traded between different counterparties but traded through a CCP can

[24] Essentially such vehicles made use of short-term funding to make long-term investments. They failed because it was simply impossible to "roll" the short-term funding.

be netted. This increases the flexibility to enter new transactions and terminate existing ones and reduces margin costs. Trading out of positions through a CCP is easy and, unlike bilateral markets, can be done with any other counterparty where the multilateral netting benefit is provided by the CCP. Furthermore, reducing the total positions that need to be replaced in the event of a default reduces price impact.

- *Loss mutualisation*. Even when a default creates losses that exceed the financial commitments from the defaulting member, these losses are distributed throughout the CCP members, reducing their impact on any one member. Thus, a counterparty's losses are dispersed partially throughout the market, making their impact less dramatic and reducing the possibility of systemic problems.
- *Legal and operational efficiency*. The margining, netting and settlement functions undertaken by a CCP potentially increase operational efficiency and reduce costs. CCPs may also reduce legal risk in providing a centralisation of rules and mechanisms. A CCP working with regulators on the best procedures is likely to be more efficient than individual market participants taking this collective responsibility.
- *Liquidity*. A CCP will improve market liquidity through the ability of market participants to trade easily and benefit from multilateral netting. Market entry is enhanced through the ability to trade anonymously and the mitigation of counterparty risk. Derivatives traded through a CCP need to be valued on a daily basis due to daily margining and cash flow payments leading to a more transparent valuation of products.
- *Transparency*. A CCP is positioned to understand the positions of all market participants and therefore is privy to potentially sensitive trading information via the overall exposure of a member. Since the CCP does not bear market risk, it has no incentive to use such information. This disperses panic that might otherwise be present in bilateral markets due to a lack of knowledge of the exposure faced by institutions. On the other hand, if a member has a particularly extreme exposure, the CCP is in a position to act on this and limit trading.
- *Default management*. A well-managed central auction is liquid and may result in smaller price disruptions than uncoordinated replacement of positions during a crisis period associated with default of a clearing member.

The distinction between bilateral and CCP-cleared OTC transactions is recognised under Basel III, which gives a relatively low 2% weighting for trade exposures to a CCP.[25]

Have CCPs Failed Before?

Historically, there are some cases of CCP failures, although they have obviously been rare. Significant examples are:

- *French Caisse de Liquidation (1973)*. This occurred as a result of a sharp drop in sugar prices and the failure of one large trading firm to post margin.
- *Commodity Clearinghouse in Kuala Lumpur (1983)*. This was a result of a crash in palm oil futures and failed margin calls from several brokers who defaulted because of the dramatic price fall.
- *Hong Kong Futures Exchange Clearing Corporation (1987)*. As described in detail by Hills *et al.* (1999), this CCP was bailed out as a result of problems with margin calls arising from the global stock market crash.

Indeed, the 1987 stock market crash, not surprisingly, caused quite severe problems for CCPs. Options traders lost large amounts of money but, since trades were reconciled only at the end of the day, many had highly leveraged positions and experienced losses far in excess of their capital. Many traders simply headed straight to the airport—the origin of the expression "airport play". This pushed US CCPs close to default as their members failed or faced severe funding strains. In the wake of the 1987 crash, both the CME and OCC had problems in receiving margin and came close to failure. Without the liquidity that the Federal Reserve injected and support from banks, a large CCP failure was more than a possibility.

Some of the lessons that can be learnt from past CCP failures and near-failures are:

- Operational risk must be controlled as much as possible (for example, after the 1987 crash, electronic reporting of trades was introduced so that the system would not be exposed to these weaknesses again).
- Variation margins should be recalculated frequently and collected promptly (intradaily if possible).

[25] Trade exposure means initial and variation margin. There is also a requirement to capitalise reserve fund exposures, which is much more complicated.

- Initial margin and reserve funds should be resilient to large negative asset shocks or gaps in market variables and to extreme co-dependency (for example, the concept of correlation increasing in a crisis).

The Impact of Homogenisation

If a major derivatives player defaults, it may not be clear how big the associated counterparty risk losses will be, nor which institutions will hear the brunt of them. This uncertainty is mitigated through a CCP allocating extreme losses across all of its members. The neutrality and ability of a CCP to disperse losses mitigates information asymmetry that can propagate stress events in bilateral markets.

A benefit of clearing is that it improves fungibility of contracts by making counterparties interchangeable and reduces systemic risk by loss mutualisation. However, this reduces the costs that riskier firms face to the detriment of the less risky firms. The former therefore can expand their trading relative to the latter. Heterogeneity of credit quality causes members' interests to diverge and they will be less likely to commit to a CCP. This could be compared to the problems that the euro currency has suffered from 2010 due to the severe credit problems of some of its member countries, especially Greece.

The homogenising of counterparty risk and use of mutualised loss-sharing reduces asymmetric informational problems and allows anonymous trading and settlement (although, as mentioned earlier, this may be costly). In a centrally cleared market using a CCP, all parties are essentially equal and the CCP acts as guarantor for all obligations. An institution has no need to assess the creditworthiness of counterparties they trade with through the CCP and may therefore reduce resources spent on monitoring individual members. They need only have confidence in the creditworthiness of the CCP. An institution with better-than-average risk management (credit quality assessment, collateral management, hedging) may lose out by trading through a CCP.

In a bilateral market, the pricing of counterparty risk will naturally cause institutions with a worsening credit quality to have higher costs and therefore provide an incentive for them to improve this aspect. However, when trading through a CCP, as long as a member is posting the relevant margin, the issue of their declining credit quality may be ignored (up to a point). This may allow poor-quality institution, to build up bigger positions more cheaply than they would normally be able to do in bilateral markets. CCPs may be more popular with counterparties with below-average risk management abilities and firms with weaker credit quality who can only achieve a limited amount of bilateral trading. The products that members may be most keen to clear through a CCP may be the more toxic ones, for example due to wrongway risk, that cannot be readily managed in a bilateral market.

Will a CCP Be Allowed to Fail?

Funnelling market activity through one institution leads tn a concentration of risk. A key component for regulators is to ensure that, especially in buoyant markets, CCPs do not become more competitive and therefore increase their likelihood of failing during volatile markets and crashes. However, despite the best efforts of CCPs, their members and regulators, one must still contemplate the possibility that, at some point, a CCP will fail.

The failure of a CCP would necessarily lead to at least a temporary breakdown of the market, as the whole structure through which positions are established, maintained and closed out would be disrupted. It would also have cross-border dimensions due to the global nature of the OTC derivatives market. Given the mandate being given to CCPs to clear many more OTC products, such a failure could be expected to be far worse than even the failure of a large bank. Whilst the *probability* of CCP failure might be smaller than that of an individual institution (thanks to tight regulation and mutualisation of losses), it represents a far more extreme and systemic event.

Relying on banks or other market participants to provide liquidity or capital to a stricken CCP may be naive, as the market conditions that caused the CCP default may mean that such institutions are also severely financially constrained. This implies that a liquidity injection from a central bank is the only way to avoid a potentially catastrophic CCP failure. Regulators are divided on whether a CCP should essentially be given such liquidity support. The obvious problem is that taxpayers bailing out a CCP is no better than bailing out other financial institutions, such as banks. Furthermore, the view that a CCP is too big to fail will lead to problems with the way CCP members (and possibly the CCP itself) behave (moral hazard again).

There is no obvious answer to the above dilemma. A critical point is that, as noted by Pirrong (2011), an institution such as a CCP could be *solvent* but *illiquid* due to massive liquidity problems in the financial markets. In such a situation, a CCP should have access to liquidity, which probably needs to come from the relevant central bank. However, this liquidity support should ideally not extend to the central bank providing an unlimited backstop for the CCP.

Could OTC Derivatives Survive without CCPs?

During the financial crisis, taxpayers essentially had to bail-out failing financial institutions to quite staggering levels and the only attempt to avoid the inherent unfairness and moral hazard from such bailouts, the Lehman bankruptcy, was an unmitigated disaster. Clearly, aspects within financial markets and OTC derivatives markets in particular need to change. Hence, it is not appropriate to argue that the move towards central clearing is simply wrong.

A reasonable question to ask might rather be whether there is a better way to mitigate systemic risk and other problems in the OTC derivatives market. In order to do this, let us recall the general functions a CCP performs:

- *Pricing and settlement*. A CCP provides the valuation of the relevant OTC derivatives and associated settlement functions.
- *Netting/trade compression*. A CCP can recognise offsetting trades. This is equivalent to netting and trade compression in bilateral markets.
- *Collateral management*. A CCP performs a collateral management function by making margin calls.
- *Reporting*. CCPs can increase transparency and provide information to regulators by maintaining records and providing reporting functionality.
- *Loss mutualisation*. A CCP provides insurance via a loss mutualisation process whereby any excess loss that is caused by the default of a CCP member is absorbed by all other CCP members.
- *Audition process*. In the event of default of a member, a CCP will auction their positions and CCP members are normally required to participate in this auction.

One could argue that of the above six tasks, bilateral markets can perform the first four adequately without CCP intervention via existing pricing/settlement methods, netting/trade compression, CSAs and trade repositories.[26] It might be that CCPs may perform them better and more efficiently but this is not completely clear as it depends on the eventual CCP landscape (for example, the number of CCPs and their product coverage). For example, considering the second point, a CCP provides multilateral netting whereas a bilateral market relies on bilateral netting and trade compression services. Only if the CCP is large enough does it reduce exposure more than other methods in the bilateral market. Furthermore, if CCPs can perform the first four functions better than bilateral markets then there would be no need to mandate clearing of any products as market participants would naturally seek CCP benefits.

The loss mutualisation point is certainly not a feature of bilateral markets. However, as discussed above, the overall benefits of this are not clear since homogeneity leads to additional problems due to moral hazard.

This leaves the final point, the auction process, as the main advantage of a CCP and the one function that is, without doubt, performed significantly better than in bilateral markets (as illustrated during the Lehman bankruptcy, for example). Any criticism of the widespread adoption of central clearing would have to provide an alternative to the CCP auction for dealing with the unwinding and/or replacement of positions resulting from a bankruptcy of a large financial institution. Specifying the protocol and legal structure for an OTC derivatives auction in the event of a major default outside CCPs is not unrealistic. The CDS market achieved something similar via the "big bang protocol" in 2009, which paved the way for auction settlement for CDS credit events, covering all CDS trades, even legacy ones that pre-dated the big bang protocol.

Hurdles and Challenges for the Growth of the CCP Market

OTC market products tend to be customised, and relatively illiquid, which limits the ability to clear them through a CCP. A certain amount of standardisation—for example, of contractual term and valuation—is required before a product can be CCP-traded. Standardisation of products to aid central clearing is a major hurdle but a necessary one, so that CCPs can offer broad product coverage. However, a danger here is that the most important products to be traded through CCPs are probably the most risky from the point of view of the stability of the CCP, and vice versa. There has been much recent interest in trading all CDS index products and single-name products through CCPs. Complexity of products may increase adverse selection problems, where dealers may have

[26] For example, the US Depository Trust & Clearing Corporation (DTCC) maintains CDS data, which it shares with authorities.

better knowledge of the inherent risk of a trade than the CCP through which they clear.

Another hurdle is cost. Moving OTC derivatives to a CCP creates costs via increased amounts and liquidity of initial margin, and a reduction in rehypothecation. It will also lead to a loss of netting benefits, although these may be recovered if the CCP market reaches a critical size.

A CCP also centralises legal and operational risks. Like all market participants, CCPs are exposed to operational risks such as systems failure and fraud. A breakdown of any aspect of a CCP's infrastructure would be catastrophic since it would affect a relatively large number of counterparties within the market. Aspects such as segregation and the movement of margin and positions through a CCP, whilst enforceable in the CCP's home jurisdiction, might be subject to legal risk from other jurisdictions.

Finally, there is the danger that the push to central clearing will influence the behaviour of market participants in potentially unintended ways. From this point of view, it is important to remain open-minded about the potential unintended consequences of CCPs and analyse their potential side-effects in as much detail as possible (e.g., see Pirrong, 2011). This does not need to be seen as a negative for central clearing but rather it can give added confidence that most of the potential weaknesses of CCPs are exposed and not hidden.

CONCLUSIONS

Bilateral OTC markets have been extremely successful and their growth has been greater than that of exchange-traded products over the last 15 years. Whilst it seems obvious that a bilaterally cleared market is more vulnerable to systemic risk, this is not an argument for the naïve introduction of CCPs. CCPs reallocate counterparty risk, but they do not make it disappear. This can be beneficial—for example, a corporate hedging their balance sheet may be wiped out if a major dealer defaults. In a CCP world, this risk is potentially reallocated to market participants who can bear it more reasonably (at a cost, of course).

CCPs reduce systemic risk (e.g., mitigating the impact of clearing member failure) and increase it (e.g., by increasing margins during a period of stress where firms may have to liquidate assets to meet large margin calls which in turn may exacerbate price volatility). Hence, CCPs transform systemic risk but do not definitely reduce it overall.

The question as to whether CCPs really reduce counterparty and/or systemic risk *overall* should be carefully considered. In bilateral markets, dealers compete for business partially based on their ability to manage counterparty risk. A CCP takes away the incentive to properly price and manage the counterparty risk created when entering a trade. Regulation may favour a certain CCP and this will create sub-optimal outcomes and market instability. A CCP would, of course, have its own risk management capabilities and be subject to prudent supervision and capital requirements in order to make its failure highly unlikely. Yet these are exactly the same measures applied to banking institutions before the 2007 crisis! The lessons from problems experienced by the Triple-A rated monolines during the 2007–2009 period must serve as a warning.

Given the global nature of OTC derivatives markets and the fact that CCPs will operate internationally, there is a need for close cross-border coordination of regulation to avoid regulatory arbitrage and mitigate systemic risk. Operational procedures should be carefully implemented and, in particular, margin should be monitored extremely carefully. CCPs clearing different markets must develop sophisticated modelling techniques to provide an aggregated assessment of the overall risk of open positions. They should perform default simulations to practice for the eventual failure of a member.

CCPs must not be focused only on clearing the products that caused the last crisis, but also those that may cause the next. The next crisis may not involve credit derivatives, indeed it may involve a class of derivatives yet to be invented. Like a financial institution, CCPs cannot be immune to financial distress. CCP access to central bank liquidity has been a controversial issue. However, CCPs could rail due to lack of liquidity. Unlimited liquidity support could lead to a bailout, of which the explicit or even implicit promise creates moral hazard. Furthermore, bailing out a CCP is ultimately no better than bailing out any other financial institution.[27]

[27] Indeed, one author has written about central clearing, "It gives the impression that regulators and legislators are reasserting control over the wild beasts of finance. In reality, the proposal may not work or materially reduce the risks it is intended to address." See "Tranquilizer solutions, Part I: A CCP idea," *Wilmott Magazine,* May 2010.

Credit Exposure

13

Learning Objectives

After completing this reading you should be able to:

- Describe and calculate the following metrics for credit exposure: expected mark-to-market, expected exposure, potential future exposure, expected positive exposure and negative exposure, effective exposure, and maximum exposure.
- Compare the characterization of credit exposure to VaR methods, and describe additional considerations used in the determination of credit exposure.
- Identify factors that affect the calculation of the credit exposure profile, and summarize the impact of collateral on exposure.
- Identify typical credit exposure profiles for various derivative contracts and combination profiles.
- Explain how payment frequencies and exercise dates affect the exposure profile of various securities.
- Explain the impact of netting on exposure, the benefit of correlation, and calculate the netting factor.
- Explain the impact of collateralization on exposure, and assess the risk associated with the remargining period, threshold, and minimum transfer amount.
- Explain the difference between risk-neutral and real-world parameters, and describe their use in assessing risk.

Excerpt is Chapter 8 of Counterparty Credit Risk and Credit Value Adjustment, *Second Edition, by Jon Gregory.*

There is no security on this earth, there is only opportunity.

General Douglas MacArthur (1880–1964)

This chapter will be concerned with defining credit exposure (often known just as exposure) in more detail and expanding the key characteristics. We then explain the important metrics used for the quantification of credit exposure. Typical exposure profiles for various instruments will be discussed and we will explain the impact of netting and collateral on exposure. Understanding exposure is relevant for the following reasons:

- Trade approval by comparing against credit limits.
- Pricing (and hedging) counterparty risk (CVA).
- Calculating economic and regulatory capital.

CREDIT EXPOSURE

Definition

A defining feature of counterparty risk arises from the asymmetry of potential losses with respect to the value[1] of the underlying transaction(s). In the event that a counterparty has defaulted, an institution may close out the relevant contract(s) and cease any future contractual payments. Following this, they may determine the net amount owing between them and their counterparty and take into account any collateral that may have been posted. Note that collateral may be held to reduce exposure but any posted collateral may have the effect of increasing exposure.

Once the above steps have been followed, there is a question of whether the net amount is positive or negative. The main defining characteristic of credit exposure is related to whether the effective value of the contracts (including collateral) is positive (in an institution's favour) or negative (against them), as illustrated in Figure 13-1.

- *Negative value.* In this case, an institution is in debt to its counterparty and is still legally obliged to settle this amount (they cannot walk away from the transaction or transactions except in specific cases). Hence, from a valuation perspective, the position appears essentially unchanged. An institution does not gain or lose from their counterparty's default in this case.
- *Positive value.* When a counterparty defaults, they will be unable to undertake future commitments and hence an institution will have a claim on the positive value at the time of the default, typically as an unsecured creditor. They will then expect to recover some fraction of their claim, just as bondholders receive some recovery on the face value of a bond. However, this unknown recovery value is, by convention, not included in the definition of exposure.

The above feature—an institution loses if the value is positive and does not gain if it is negative—is a defining characteristic of counterparty risk. We can define exposure simply as

$$\text{Exposure} = \text{Max}(\text{value}, 0) \quad \textbf{(13.1)}$$

This would mean that defining exposure at a given time is rather easy. One simply values the relevant contracts, aggregates them according to the relevant netting rules and finally applies an adjustment corresponding to any collateral held against the positions. This appears to be a reasonably simple measure of exposure and is usually the one applied in pricing and risk management calculations.

Bilateral Exposure

A key feature of counterparty risk is that it is bilateral as both parties to a transaction can default and therefore both can experience losses. For completeness, one must consider losses arising from both defaults. From an institution's point of view, their own default will cause a loss to any counterparty they are in debt to. This can be defined in terms of *negative exposure*, which by symmetry is

$$\text{Negative exposure} = \text{Min}(\text{value}, 0) \quad \textbf{(13.2)}$$

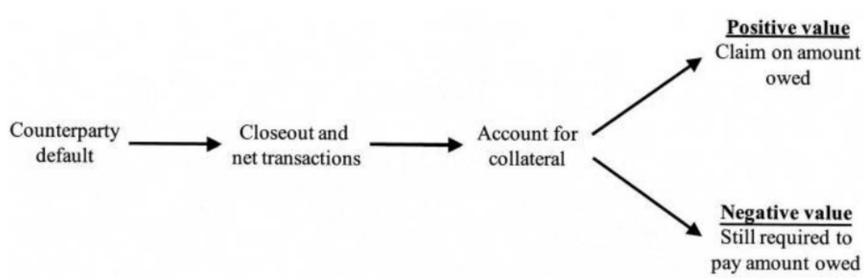


FIGURE 13-1 Illustration of the impact of a positive or negative value in the event of the default of a counterparty.

[1] The definitions of value will be more clearly discussed below.

TABLE 13-1 Illustration of Payoff in the Event of an Institution or Their Counterparty Defaulting

Rec_C and Rec_I denote recovery values for the counterparty and institution respectively.

Event	Exposure	Impact	Payoff
Counterparty defaults	Exposure	Loss	$Rec_C \times$ Max(value, 0)
Institution defaults	Negative exposure	Gain	$-Rec_I \times$ Min(value, 0)

A negative exposure leads to a gain, which is relevant since the counterparty is making a loss.[2]

Finally, it is important to note that recovery values are relevant in the case of exposure and negative exposure. In the former case, an institution will be paid only a recovery fraction (Rec_C) of their exposure in the event their counterparty defaults. This is a loss for the institution. In the latter case, the counterparty will receive only a recovery fraction (Rec_I) of the negative exposure in the event the institution defaults. This is a gain for the institution. The bilateral payoffs defining counterparty risk are illustrated in Table 13-1.

The Closeout Amount

The amount represented by "value" above represents the valuation of the relevant contracts at the default time of the counterparty (or institution), including the impact of netting and collateral. An institution therefore would presumably aim for the relevant documentation and legal practices to align this amount as closely as possible to the value of the contracts (and collateral) from their point of view. In practice, this is not completely trivial as it is necessary to agree with the counterparty (or the administrators of their default) the relevant valuations. In cases where such an agreement cannot be reached, the valuation must be made by following the relevant procedures in the documentation.

ISDA documentation uses the term "closeout amount" to define what is referred to above as "value." However, the precise definition of closeout is, perhaps not unreasonably, rather vague[3] and can give rise to several issues. In determining a closeout amount, according to ISDA (2009), an institution should "act in good faith and use commercially reasonable procedures in order to produce a commercially reasonable result." Whilst efforts have been made to define closeout language in the most appropriate way and to learn from problems relating to historic bankruptcies, there will clearly always be the chance of a dispute over closeout amounts.

Whilst subjectivities and disagreements over the precise definition of the closeout amount may lead to uncertainties, this should still correspond closely to a standard risk-free value appearing in Equation (13.1) corrected for any relevant costs in replacing transactions. This would make the above simple definition of exposure a reasonable proxy, albeit with some (hopefully small) uncertainty due to the difficulty in defining the closeout amount. However, there is also a potentially important and systematic effect. ISDA (2009) specifies that the closeout amount may include information related to the creditworthiness of the surviving party. This implies that an institution can potentially reduce the amount owed to a defaulting counterparty or increase their claims in accordance with their *own* counterparty risk, which would be relevant in a replacement transaction. In the event the institution defaults, then the counterparty is able to do the same. The result of this is a recursive problem where the very definition of an institution's counterparty risk depends on their own counterparty risk in the future. We will base exposure on standard risk-free valuation.

A final point to note about the above problems in determining closeout amounts is the time delay. Until an agreement is reached, an institution cannot be sure of the precise amount owed or the value of their claim as an unsecured creditor. This will create particular problems

[2] This is a symmetry effect where one party's gain must be another's loss. There may be reasonable concern with defining a gain in the event of an institution's own default.

[3] Historically, there have been two methods to govern the calculation of closeout amounts. The first is a "Market Quotation" approach that requires quotes from dealers (a minimum number of quotes may be specified) in the relevant market for replacement transactions that represent the economic equivalent of those being terminated. However, in cases of extreme volatility and even a breakdown of the market (following a major default, for example) it may still be difficult to achieve the quote required to establish such an amount. Another approach is the "Loss" method, where an institution may determine its losses (including costs) because of terminating the relevant transactions and re-establishing the equivalent positions.

for managing counterparty risk. In a default involving many contracts (such as the number of OTC derivatives and the Lehman bankruptcy), the sheer operational volumes can make the time taken to agree on such valuations considerable.

Exposure as a Short Option Position

Counterparty risk creates an asymmetric risk profile as shown by Equation (13.1). When a counterparty defaults, an institution loses if the value is positive but does not gain if it is negative. The profile can be likened to a short[4] option position. Familiarity with basic options-pricing theory would lead to two obvious conclusions about the quantification of exposure:

- Since exposure is similar to an option payoff, a key aspect will be *volatility* (of the value of the relevant contracts and collateral).
- Options are relatively complex to price (compared with the underlying instruments at least). Hence, to quantify credit exposure even for a simple instrument may be quite complex.

By symmetry, an institution has long optionality from their own default. We note that this is not relevant from a risk management or regulatory point of view but may be relevant for pricing.

Future Exposure

A current valuation of all relevant positions and collateral will lead us to a calculation of current exposure (admittedly with some uncertainty regarding the actual closeout amount). However, it is even more important to characterise what the exposure might be at some point in the future. This concept is illustrated in Figure 13-2, which can be considered to represent any situation from a single trade to multiple netted trades with some collateral arrangement associated with them. Whilst the current (and past) exposure is known with certainty, the future exposure is defined probabilistically by what may happen in the future in terms of market movements and contractual features of transactions, both of which are uncertain. Hence, in understanding future exposure one must define the level of the exposure and also its underlying uncertainty.

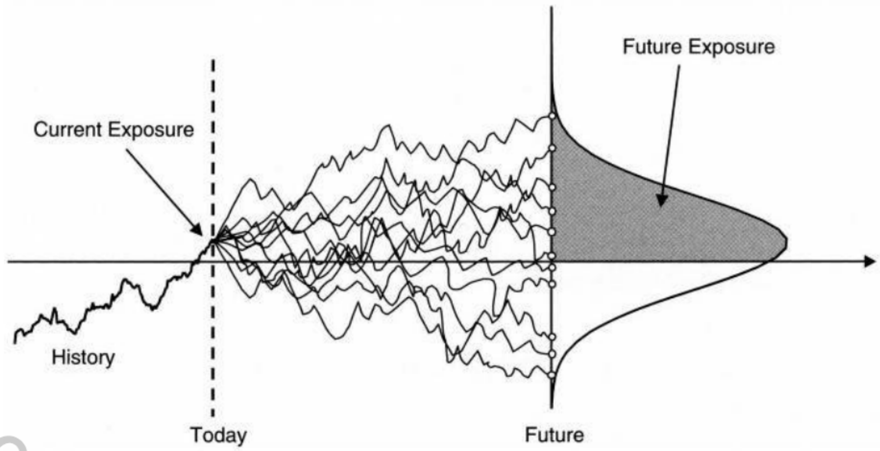


FIGURE 13-2 Illustration of future exposure with the grey area representing exposure (positive future values). The white area represents negative exposure.

Quantifying exposure is extremely complex due to the long periods involved, the many different market variables that may influence the exposure, and risk mitigants such as netting and collateral. This chapter focuses on defining exposure and discussing intuitively the impact of aspects such as netting and collateral.

Comparison to Value-at-Risk

In financial risk management VaR methods have, for almost two decades, been a popular methodology to characterise market risk. Any reader familiar with VaR will recognise from Figure 13-2 that the characterisation of exposure shares similarities with the characterisation of VaR. This is indeed true, although we note that in characterising exposure we are faced with additional complexities, most notably:

- *Time horizon.* Unlike VaR, exposure needs to be defined over multiple time horizons (often far in the future) so as to understand fully the impact of time and specifics of the underlying contracts. There are two important implications of this. The first is that "ageing" of transactions must be considered. This refers to understanding a transaction in terms of all future contractual payments and changes such as cash flows, exercise

[4] The short option position arises since exposure constitutes a loss.

decisions, cancellations (e.g., callability) and more exotic aspects such as barrier crossings. Such effects may also create path dependency where the exposure at one date depends on an event defined at a previous date. In VaR models, due to the 10-day horizon used,[5] such aspects can be neglected. The second important point here is that, when looking at longer time horizons, the trend (also known as drift) of market variables (in addition to their underlying volatility and co-dependence structure) is relevant (as depicted in Figure 13-2). In VaR the drift can be ignored, again, since the period is so short.

- *Risk mitigants.* Exposure is typically reduced by risk mitigants such as netting and collateral and the impact of these mitigants must be considered in order to properly estimate future exposure. In some cases, such as applying the correct netting rules, this requires knowledge of the relevant legal agreements. However, in the case of future collateral amounts, another degree of subjectivity is created since there is no certainty over the type of collateral and precise time that it would be received. Other contractual features of transactions, such as termination agreements, may also create subjectivity and all such elements must be modelled, introducing another layer of complexity and uncertainty.
- *Application.* VaR is a risk management approach. Exposure must be defined for both risk management and pricing (i.e., CVA). This creates additional complexity in quantifying exposure and may lead to two completely different sets of calculations, one to define exposure for risk management purposes and one for pricing purposes (Chapter 15).

In other words, exposure is much more complex than VaR (which is itself a complex concept). Moreover, exposure is only one component of counterparty risk. However, no one told you this was going to be easy, did they?

METRICS FOR CREDIT EXPOSURE

In this section, we define the measures commonly used to quantify exposure. The different metrics introduced will be appropriate for different applications. There is no standard nomenclature used and some terms may be used in other contexts elsewhere. We follow the Basel Committee on Banking Supervision (2005) definitions, which are probably the most commonly used although, unfortunately, not the most intuitively named.

We begin by defining exposure metrics for a given time horizon. Note that in discussing exposure below, we are referring to the total number of relevant trades, netted appropriately and including any relevant collateral amounts. We will refer to this as the netting set.

Expected Future Value

This component represents the forward or expected value of the netting set at some point in the future. As mentioned above, due to the relatively long time horizons involved in measuring counterparty risk, the expected value can be an important component, whereas for market risk VaR assessment (involving only a time horizon of 10 days), it is not. Expected future value (EFV) may vary significantly from current value for a number of reasons:

- *Cash flow differential.* Cash flows in derivatives transactions may be rather asymmetric. For example, early in the lifetime of an interest rate swap, the fixed cash flows will typically exceed the floating ones (assuming the underlying yield curve is upwards sloping as is most common). Another example is a cross-currency swap where the payments may differ by several percent annually due to a differential between the associated interest rates. The result of asymmetric cash flows is that an institution may expect a transaction in the future to have a value significantly above or below the current one. Note that this can also apply to transactions maturing due to final payments.
- *Forward rates.* Forward rates can differ significantly from current spot variables. This difference introduces an implied drift (trend) in the future evolution of the underlying variables in question (assuming one believes this is the correct drift to use). Drifts in market variables will lead to a higher or lower future value for a given netting set even before the impact of volatility. Note that this point (although not always) may be simply another way to state the above point on cash flow differential since in many transactions such as swaps, the

[5] In many cases this is a 1-day horizon, which is simply scaled to 10 days.

cash flow differential is essentially a result of forward rates being different from spot rates.

- *Asymmetric collateral agreements.* If collateral agreements are asymmetric (such as a one-way CSA) then the future value may be expected to be higher or lower reflecting respectively unfavourable or favourable collateral terms.

Potential Future Exposure

In risk management, it is natural to ask ourselves *what is the worse exposure we could have at a certain time in the future.* PFE will answer this question with reference to a certain confidence level. For example, the PFE at a confidence level of 99% will define an exposure that would be exceeded with a probability of no more than 1% (one minus the confidence level). PFE is therefore exactly the same measure as VaR except for the obvious differences that PFE needs to be defined at more than one future date and represents gains[6] (exposure) rather than losses. PFE is illustrated in Figure 13-3. Note[7] also that, as shown, the centre of the distribution can differ significantly from zero (this represents the future value of the transactions having a significantly positive or negative expected value).

The confidence level is consistent with the probability of losses exceeding the PFE. For example, a 99% confidence level means that the small probability of a loss greater than PFE in Figure 13-3 is 1%.

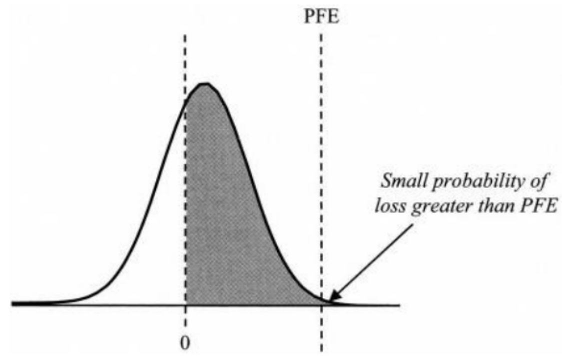


FIGURE 13-3 Illustration of potential future exposure.

The grey area represents positive values which are exposures.

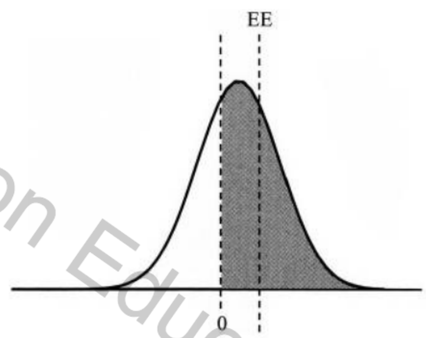


FIGURE 13-4 Illustration of expected exposure.

The grey area represents positive values, which represent exposures.

Expected Exposure

In addition to PFE, which is clearly a risk management measure, the pricing of counterparty risk will involve expected exposure (EE), illustrated in Figure 13-4. This is the average of all exposure values. Note that only positive values (the grey area) give rise to exposures and other values have a zero contribution (although they contribute in terms of their *probability*). This means that the expected exposure will be above the expected future value (this is similar to the concept of an option being more valuable than the underlying forward contract).

[6] Such a gain may not be an actual gain as there will be a corresponding hedge with a different counterparty.

[7] Note that the normal distribution used to depict the distribution of future values does not need to be assumed and also that PFE is often defined at confidence levels other than 99%.

EE and PFE for a Normal Distribution

In Appendix A we give simple formulas for the EE and PFE for a normal distribution. These formulas are reasonably simple to compute and will be useful for some examples used throughout this book.

SPREADSHEET 13-1 EE and PFE for a Normal Distribution

To download Spreadsheet 13-1, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 8 exercises.

Example

Suppose future value is defined by a normal distribution with mean 2.0 and standard deviation 2.0. As given by the formulas in Appendix A, the EE and PFE (at the 99% confidence level) are

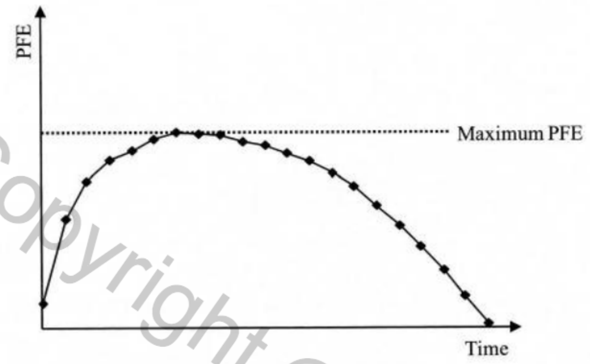


FIGURE 13-5 Illustration of maximum PFE.

$$\text{EE} = 2.17$$
$$\text{PFE} = 6.65$$

If the standard deviation was increased to 4.0, we would obtain

$$\text{EE} = 2.79$$
$$\text{PFE} = 11.31$$

Note that the EE, like the PFE, is sensitive to standard deviation (volatility).

Maximum PFE

Maximum or peak PFE simply represents the highest PFE value over a given time interval, thus representing the worst-case exposure over the entire interval. This is illustrated in Figure 13-5.

Expected Positive Exposure

Since EE is already an average over all exposures, it is perhaps natural to continue this averaging over time. Expected positive exposure (EPE) is defined as the average exposure across all time horizons. It can therefore

SPREADSHEET 13-2 EPE Example

To download Spreadsheet 13-2, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 8 exercises.

be represented as the weighted average of the expected exposure across time, as illustrated in Figure 13-6.

This single EPE number is often called a "loan equivalent," as the average exposure is equivalent to the average amount lent (via an exposure) to the counterparty in question. It is probably obvious that expressing a highly

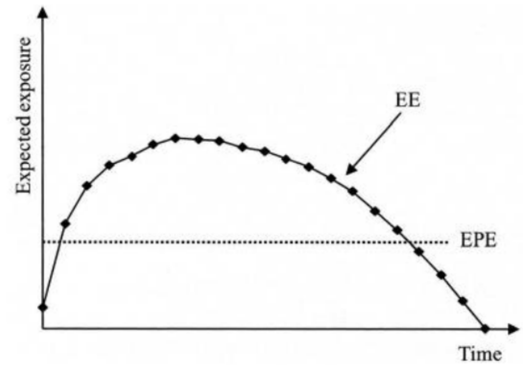


FIGURE 13-6 Illustration of expected positive exposure, which is the weighted average (the weights being the time intervals) of the EE profile.

uncertain exposure by a single EPE or loan-equivalent amount can represent a fairly crude approximation, as it averages out both the randomness of market variables and the impact of time. However, we shall see later that EPE has a strong theoretical basis for pricing (Chapter 15) and assessing portfolio counterparty risk.

Negative Exposure

Exposure is represented by positive future values. Conversely, we may define negative exposure as being represented by negative future values. This will obviously represent the exposure from a counterparty's point of view. We can therefore define measures such as negative expected exposure (NEE) and expected negative exposure (ENE), which are the precise opposite of EE and EPE.

Effective Expected Positive Exposure

Measures such as EE and EPE may underestimate exposure for short-dated transactions and not capture properly "roll-over risk." This arises from current short-dated transactions that will be rolled over into new transactions at their maturity. For these reasons, the terms effective EE and effective EPE (EEPE) were introduced by the Basel Committee on Banking Supervision (2005). Effective EE is simply a non-decreasing EE. Effective EPE is the average of the effective EE. These terms are shown in comparison with EE and EPE in Figure 13-7. These measures approximately assume that any maturing transactions will be replaced.[8] The role and definition of EEPE is discussed in more detail in Chapter 14.

[8] They essentially assume that any reduction in exposure is a result of maturing transactions. This is not necessarily the case.

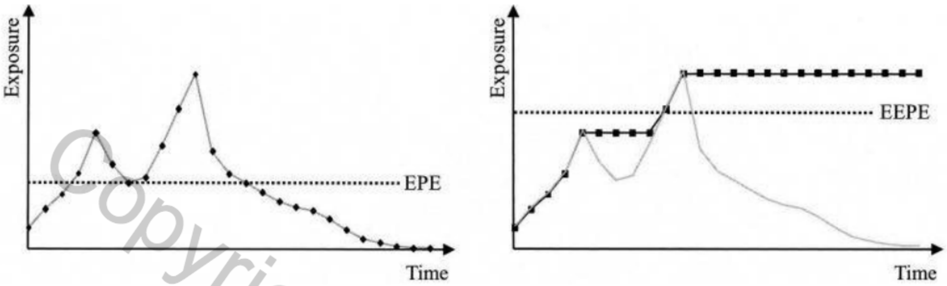


FIGURE 13-7 Illustration of effective EE and effective EPE.

SPREADSHEET 13-3 EPE and Effective EPE Example

To download Spreadsheet 13-3, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 8 exercises.

We emphasise that some of the exposure metrics defined above, whilst common definitions, are not always used. The definitions above are used by the Basel Committee on Banking Supervision (2005).

FACTORS DRIVING CREDIT EXPOSURE

We now give some examples of the significant factors that drive exposure, illustrating some important effects such as maturity, payment frequencies, option exercise, roll-off and default. Our aim here is to describe some key features that must be captured. In all the examples below, we will depict PFE defined as a percentage of the notional of the transaction in question.

Loans and Bonds

The exposures of debt instruments such as loans and bonds can usually be considered almost deterministic and approximately equal to the notional value. Bonds typically pay a fixed rate and therefore will have some additional uncertainty since, if interest rates decline, the exposure may increase and vice versa. In the case of loans, they are typically floating-rate instruments but the exposure may decline over time due to the possibility of prepayments.

Future Uncertainty

The first and most obvious driving factor in exposure is future uncertainty. Forward contracts such as forward rate agreements (FRAs) and FX forwards are usually characterised by having just the exchange of two cash flows or underlyings (often netted into a single payment) at a single date, which is the maturity of the contract. This means that the exposure is a rather simple increasing function reflecting the fact that, as time passes, there is increasing uncertainty about the value of the final exchange. Based on normal distribution assumptions, such a profile will follow a "square root of time" rule, meaning that it will be proportional to some constant times the square root of the time (t):

$$\text{Exposure} \propto \sqrt{t} \qquad (13.3)$$

Such a profile is illustrated in Figure 13-8. We can see from the above formula that the maturity of the contract does not influence the exposure (except for the obvious reason that there is zero exposure after this date). For similar reasons, a similar shape is seen for vanilla options with an upfront premium. More exotic options may have more complex profiles.

Periodic Cash Flows

Many OTC derivatives include the periodic payment of cash flows, which has the impact of reversing the effect of future uncertainty. The most obvious and common example here is a swap which is characterised by a peaked

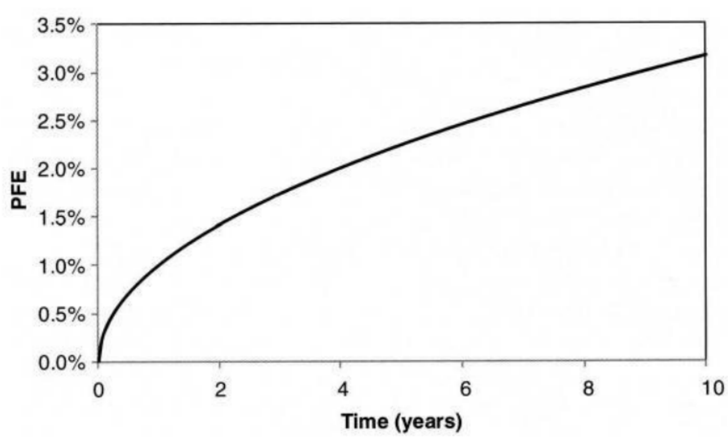


FIGURE 13-8 Illustration of a square root of time exposure profile that is seen, for example, in forward contracts and vanilla options positions.

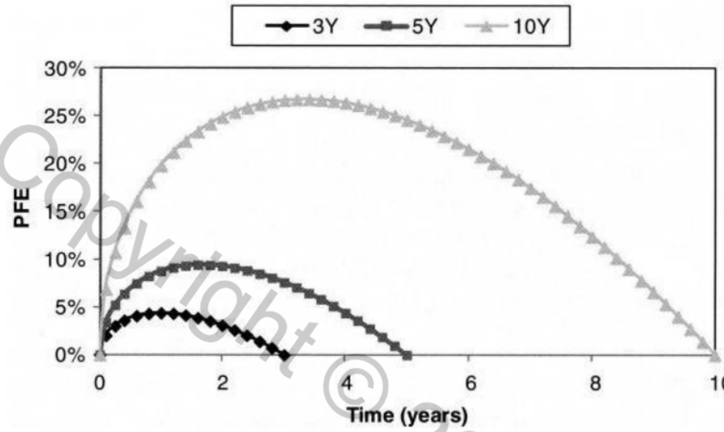


FIGURE 13-9 Illustration of the PFE of swaps of different maturities.

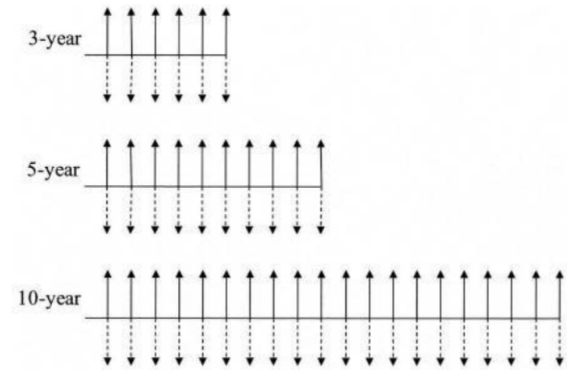


FIGURE 13-10 Illustration of a cash flow swap transaction of different maturities.

shape as shown in Figure 13-9. The shape arises from the balance between future uncertainties over payments, combined with the roll-off of swap payments over time. This can be represented approximately as

$$\text{Exposure} \propto (T - t)\sqrt{t} \qquad \textbf{(13.4)}$$

where T represents the maturity of the trade in question. The above function is initially increasing due to the $\sqrt{t}$ term but then decreases to zero as a result of the $(T - t)$ component, which is an approximate representation of the remaining maturity of the trade at a future time t. It can be shown that the maximum of the above function occurs at $T/3$, i.e., the maximum exposure occurs at one-third of the lifetime.

A swap with a longer maturity has much more risk due to both the increased lifetime and the greater number of payments due to be exchanged. An illustration of the swap cash flows is shown in Figure 13-10.

An exposure profile can be substantially altered due to the more specific nature of the cash flows in a transaction. A basis swap where the payments are made more frequently than they are received will then have more risk than the equivalent equal payment swap. This effect is illustrated in Figures 13-11 and 13-12.

Another impact the cash flows have on exposure is in creating an asymmetry between opposite trades. In the case of an interest rate swap, this occurs because of the different cash flows being exchanged. In a "payer swap," fixed cash flows are paid periodically at a deterministic amount (the "swap rate") whilst floating cash flows are received.

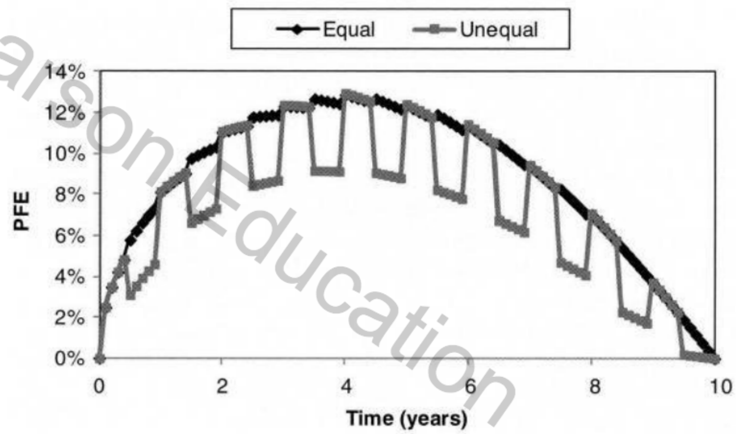


FIGURE 13-11 Illustration of PFE for swaps with equal and unequal payment frequencies.

The latter corresponds to a swap where cash flows are received quarterly but paid only semi-annually.

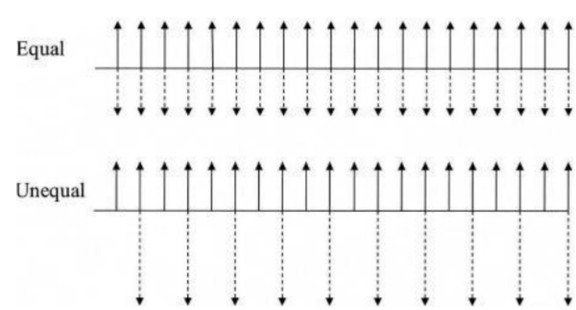


FIGURE 13-12 Illustration of the cash flows in swap transactions with different payment frequencies.

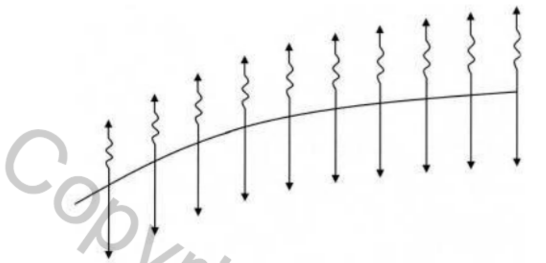


FIGURE 13-13 Illustration of the floating against fixed cash flows in a swap where the yield curve is upwards sloping.

Whilst the risk-neutral expected value of the floating and fixed cash flows is equal, the projected floating cash flows are expected to be smaller at the beginning and larger at the end of the swap.

The value of future floating cash flows is not known until the fixing date although, at inception, their risk-neutral value will be equal to that of the fixed cash flows. The value of the projected[9] floating cash flows depends on the shape of the underlying yield curve. In the case of a typical upwards-sloping yield curve, the initial floating cash flows will be expected to be smaller than the fixed rate paid whilst later in the swap the trend is expected to reverse. This is illustrated schematically in Figure 13-13.

The net result of this effect is that the exposure of the payer swap is higher due to the expectation to pay net cash flows (the fixed rate against the lower floating rate) in the first periods of the swap and receive net cash flows later in the lifetime (Figure 13-14). Another way to state this is that the forward value of the swap is positive (by an amount defined by the expected net cash flows). The opposite "receiver" swap has a correspondingly lower exposure.

The above effect can be even more dramatic in cross-currency swaps where a high-interest-rate currency is paid against one with lower interest rates (as was the case, for example, with widely traded dollar versus yen swaps for many years before the dramatic US interest rate cuts of 2008/09), as illustrated in Figure 13-15. The overall high interest rates paid are expected to be offset by the gain on the notional exchange at the maturity of the contract,[10] and this expected gain on exchange of

[9] By projected we mean the risk-neutral expected value of each cash flow.

[10] From a risk-neutral point of view.

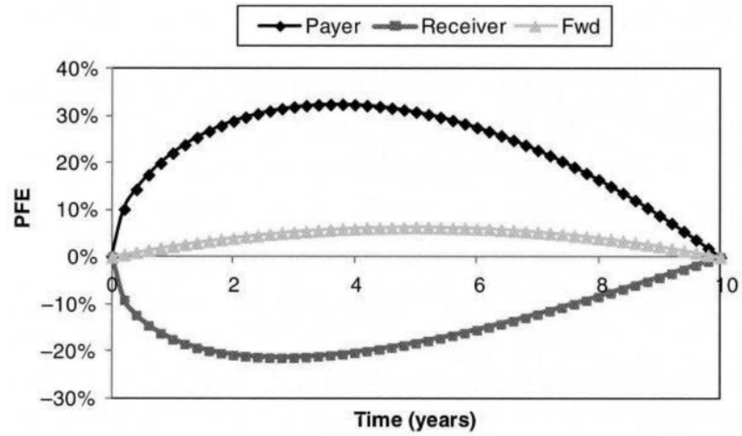


FIGURE 13-14 Illustration of the PFE for payer and receiver swaps and the associated forward value.

The receiver swap is shown for ease of exposition as a negative exposure.

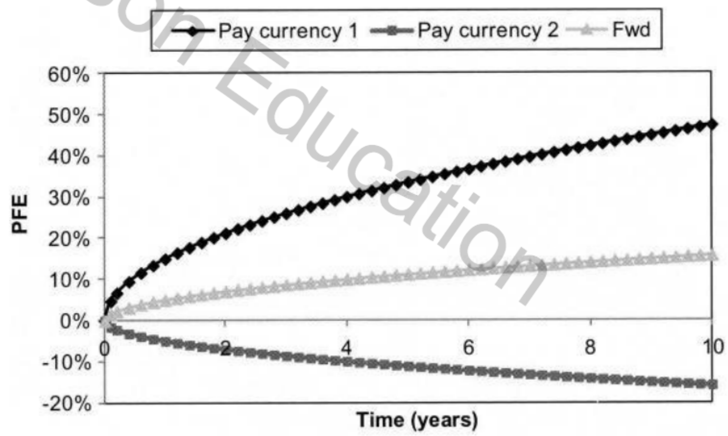


FIGURE 13-15 Illustration of the PFE for cross-currency swaps with the associated forward value with currency 1 having higher interest rates than currency 2.

The latter case is shown for ease of exposition as a negative exposure.

notional leads to a significant exposure for the payer of the high interest rate. In the reverse swap, it is increasingly likely that there will be a negative MtM on the swap when paying the currency with the lower interest rates. This creates a "negative drift," making the exposure much lower.

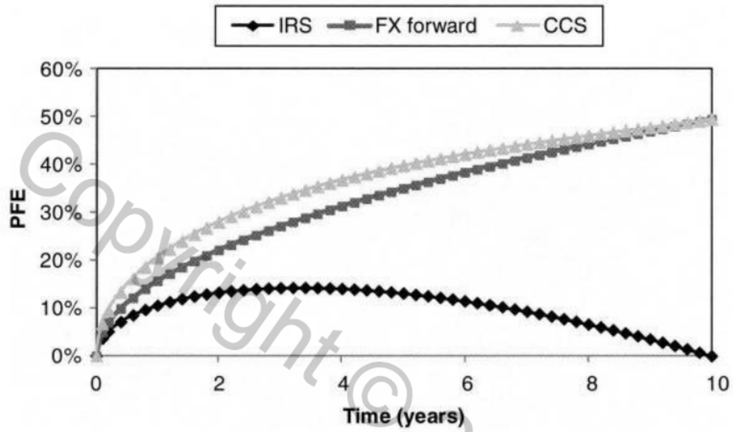


FIGURE 13-16 Illustration of a cross-currency swap (CCS) profile as a combination of an interest rate swap (IRS) and FX forward.

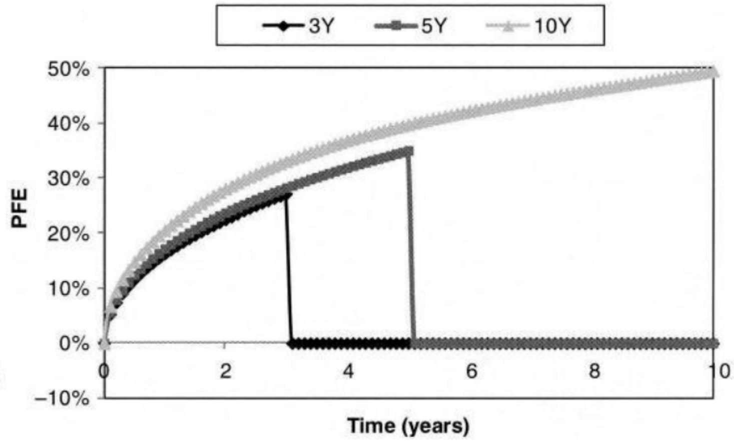


FIGURE 13-17 Illustration of the PFE of cross-currency swaps of different maturities.

Combination of Profiles

Some products have an exposure that is driven by a combination of two or more underlying risk factors. An obvious example is a cross-currency swap, which is essentially a combination of an interest rate swap and an FX forward trade.[11] This would therefore be represented by a combination of the profiles shown in Figures 13-8 and 13-9. Figure 13-16 illustrates the combination of two such profiles. Foreign exchange exposures can be considerable due to the high FX volatility driving the risk coupled to the long maturities and final exchanges of notional. The contribution of the interest rate swap is typically smaller, as shown. We note also that the correlation between the two interest rates and the FX rate is an important driver of the exposure (in Figure 13-16 a correlation of 20% is assumed, which increases the cross-currency exposure[12]).

SPREADSHEET 13-4 Simple CCS Profile

To download Spreadsheet 13-4, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 8 exercises.

In Figure 13-17 we illustrate the exposure for cross-currency swaps of different maturities. The longer-maturity swaps have marginally more risk due to the greater number of interest rate payments on the swap.

Optionality

The impact of exercise decisions creates some complexities in exposure profiles. In Figure 13-18, we show the exposure for a European-style interest rate swaption that is swap-settled (physical delivery) rather than cash-settled.[13] The underlying swap has different payment frequencies also. We compare it with the equivalent forward swap. Before the exercise point, the swaption must always have a greater exposure than the forward swap[14] but thereafter, this trend will reverse since there will be scenarios where the forward swap has positive value but the swaption would not have been exercised. This effect is illustrated in Figure 13-19, which shows a scenario that would give rise to exposure in the forward swap but not the swaption.

[11] Due to the exchange of notional at the end of the transaction.

[12] The impact of correlation can be seen in Spreadsheet 13-4. We note that the correlation here is not directly between the two interest rates and the FX rate, as would be required in practice.

[13] The cash-settled swaption has an identical exposure until the exercise date and then zero exposure thereafter. Physically settled swaptions are standard in some interest rate markets.

[14] The option to enter into a contract cannot be worth less than the equivalent obligation to enter into the same contract.

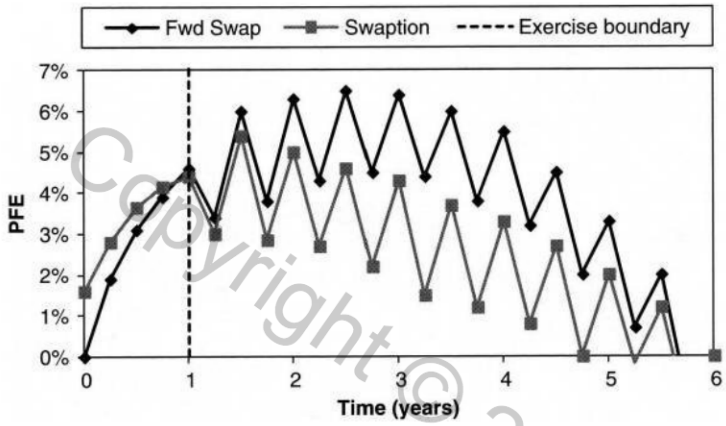


FIGURE 13-18 PFE for a swap-settled (physically settled) interest rate swaption and the equivalent forward swap.

The option maturity is 1 year and the swap maturity 5 years.

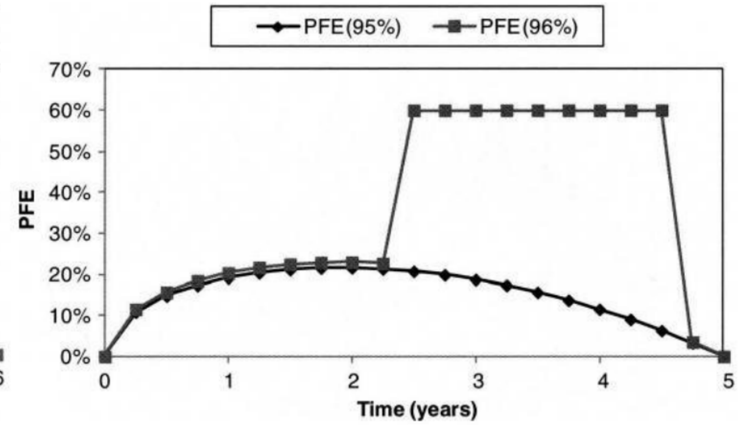


FIGURE 13-20 PFE for a long protection single-name CDS trade computed at confidence levels of 95% and 96%.

A PFE of 60% arises from default with an assumed recovery rate of 40%.

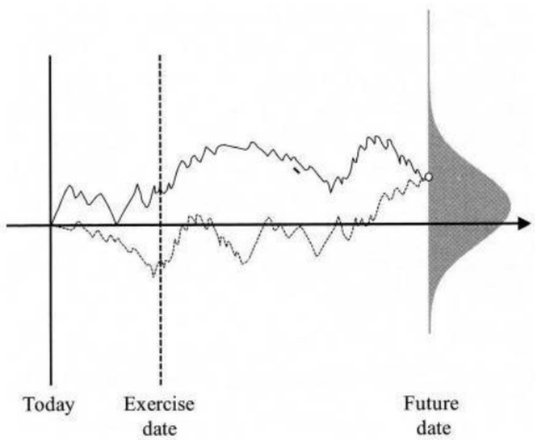


FIGURE 13-19 Illustration of exercise of a physically settled European swaption showing two potential scenarios of future value for the underlying swap.

The solid line corresponds to a scenario where the swaption would be exercised, giving rise to an exposure at the future date. The dotted line shows a scenario that would give rise to an identical exposure but where the swaption would not have been exercised. The exercise boundary (point at which the swaption is exercised) is assumed to be the *x*-axis.

We can make a final comment about the swaption example, which is that surely in exercising one should incorporate the views on counterparty risk at that time. In other words, CVA should be a component in deciding whether to exercise or not. Yet we have not calculated CVA yet! We will discuss this recursive problem again in Chapter 15.

Credit Derivatives

Credit derivatives represent a big problem for counterparty risk assessment due to wrong-way risk, which will be discussed extensively in Chapter 16. Even without this as a consideration, exposure profiles of credit derivatives are hard to characterise due to the discrete payoffs of the instruments. Consider the exposure profile of a single-name CDS as shown in Figure 13-20 (long CDS protection). The exposure increases in the early stages, which corresponds to scenarios in which the CDS premium (credit spread) will have widened. However, the maximum exposure on the CDS corresponds to the reference entity experiencing a credit event, which triggers an immediate payment of the notional less a recovery value (60% in the example, assuming a 40% recovery value). This is a rather unnatural effect[15] (see also Hille et. al., 2005), as it means that PFE may or may not represent the actual credit event occurring and is sensitive to the confidence level used. In the example, at 3 years the 95% PFE is defined by a large credit spread widening whilst the 96% PFE is defined by the credit event (and is three

[15] We comment that the above impact could be argued to be largely a facet of common modelling assumptions, which assume default as a sudden unanticipated jump event with a known recovery value (40%). Using a more realistic modelling of default and an unknown recovery value gives behaviour that is more continuous.

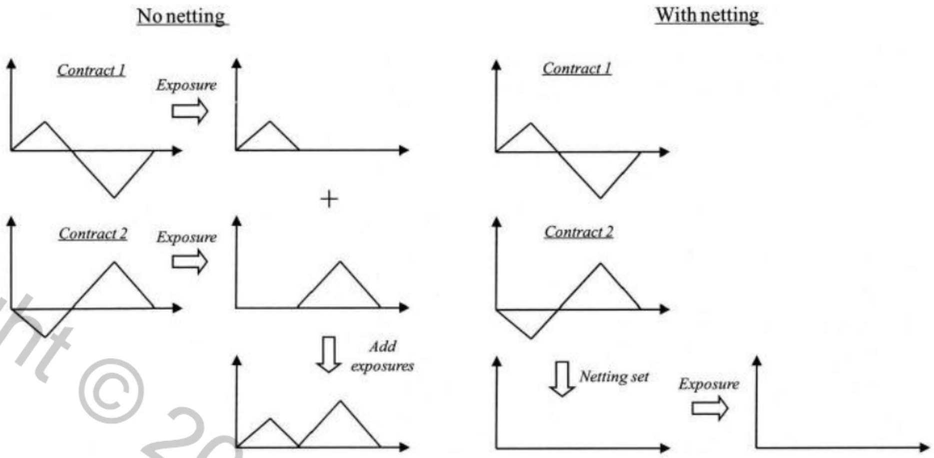


FIGURE 13-21 Illustration of the impact of netting on exposure.

times larger). Using a measure such as expected shortfall[16] partially solves this problem.

SPREADSHEET 13-5 CDS PFE Example

To download Spreadsheet 13-5, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 8 exercises.

UNDERSTANDING THE IMPACT OF NETTING ON EXPOSURE

We now consider in more depth the impact of netting on exposure. Since netting allows the future values of trades to offset one another, then the aggregate effect of all trades must be considered. As we shall see, there are several different aspects to contemplate before understanding the full netting impact on overall exposure with respect to a particular counterparty. In the next section, we will consider the impact of collateral on exposure.

[16] Expected shortfall is a measure used in preference to VaR in some cases since it has more mathematically convenient properties and, unlike VaR, is a "coherent risk measure." In this case, it corresponds to the expected exposure conditional on being above the relevant PFE value.

The Impact of Netting on Future Exposure

We illustrate the impact of netting on exposure in Figure 13-21 with exactly opposite transactions. When there is no legal agreement to allow netting, exposures must be considered additive. This means that the positions do not offset one another. With netting allowable (and enforceable), one can add values at the netting set level before calculating the exposure and therefore the profiles shown give a zero exposure at all points in the future.

Netting and the Impact of Correlation

When considering the netting benefit of two or more trades, the most obvious consideration is the correlation between the future values (and therefore exposures also). A high positive correlation between two trades means that future values are likely to be of the same sign. This means that the netting benefit will be small or even zero. We illustrate this in Table 13-2, where we can see that the two sets of values create very little netting benefit. Netting will only help in cases where the values of the trades have opposite signs, which occurs only in scenario 3. The EE (average of the exposures assuming equally weighted scenarios) is reduced by only a small amount.

Highly correlated exposures[17] (as in Table 13-2) will provide the least netting benefit and, in case of identical

[17] Note that the future values in Table 13-2 have a correlation of 100% but the exposures do not.

TABLE 13-2 Illustration of the Impact of Netting When There Is Positive Correlation between MtM Values. The expected exposure is shown assuming each scenario has equal weight.

	Future Value		Total Exposure		
	Trade 1	**Trade 2**	**No Netting**	**Netting**	**Netting Benefit**
Scenario 1	25	15	40	40	0
Scenario 2	15	5	20	20	0
Scenario 3	5	−5	5	0	5
Scenario 4	−5	−15	0	0	0
Scenario 5	−15	−25	0	0	0
EE			13	12	1

TABLE 13-3 Illustration of the Impact of Netting When There Is Negative Correlation between Future Values. The expected exposure is shown assuming each scenario has equal weight.

	Future Value		Total Exposure		
	Trade 1	**Trade 2**	**No Netting**	**Netting**	**Netting Benefit**
Scenario 1	25	−15	25	10	15
Scenario 2	15	−5	15	10	5
Scenario 3	5	5	10	10	0
Scenario 4	−5	15	15	10	5
Scenario 5	−15	25	25	10	15
EE			18	10	8

distributions (add 10 to each scenario for trade 2 to see this effect), this simply corresponds to increasing the size of a given transaction, in which case there will be no netting benefit at all. Negative correlations are clearly much more helpful as future values are much more likely to have opposite signs and hence the netting benefit will be stronger. We illustrate this in Table 13-3, where we see that netting is beneficial in four out of the five scenarios. The EE is almost half the value without netting.

The extreme case of perfect negative correlation (as in Table 13-3) will provide the maximum netting benefit. In the case of identical distributions (subtract 10 from each scenario for trade 2 to see this effect), this simply corresponds to perfectly offsetting transactions (perhaps due to a cancellation via an unwind), in which case the netting benefit is 100% since there is no overall risk.

A majority of netting may occur across instruments of different asset classes that may be considered to have only a small correlation. One should note that this would still create a positive benefit. Indeed, for a simple example in Appendix B, we show the reduction corresponding to the case of normal variables with zero mean and equal variance. We derive the following formula for the "netting factor" with respect to exposure under the assumption that future values follow a multivariate normal distribution:

$$\text{Netting factor} = \frac{\sqrt{n + n(n-1)\bar{\rho}}}{n} \quad \textbf{(13.5)}$$

where n represents the number of exposures and $\bar{\rho}$ is the average correlation. The netting factor represents the ratio of net to gross exposure and will be +100% if there is no netting benefit ($\bar{\rho}$ = 100%) and 0% if the netting

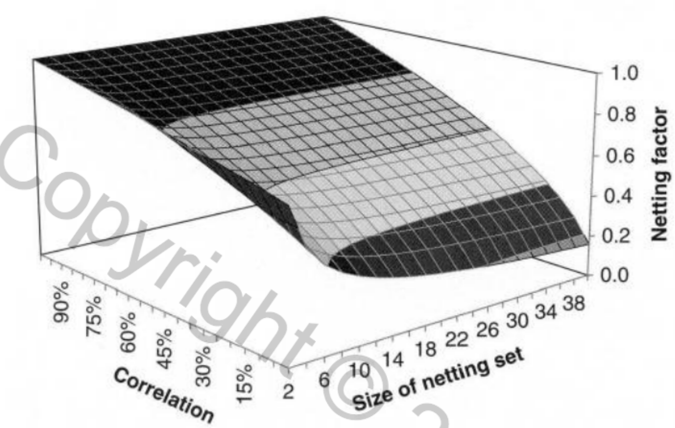


FIGURE 13-22 Illustration of the netting benefit in a simple example as a function of the size of the netting set (number of trades) and correlation as derived in Appendix B.

Only positive correlations are shown.

benefit is maximum ($\bar{\rho} = -(n-1)^{-1}$)[18]. We illustrate the above expression in Figure 13-22, where we can see that the netting benefit improves (lower netting value) for a large number of exposures and low correlation as one would expect, since these conditions maximise the diversification benefit. We note that this is a stylised example but it shows the general impact of correlation and the size of the netting set.

[18] Note that there is a restriction on the correlation level that ensures the term inside the square root in Equation (13.5) does not become negative. This is explained in Appendix B.

With no correlation, the simple formula tells us that the overall netting factor for n exposures is $1/\sqrt{n}$. This means, for example, that two independent exposures with zero mean and equal volatility have a netted exposure reduced to 71% of the exposure without netting. For five exposures, the netting factor decreases to 45%.

Netting and Absolute Value

In Table 13-2, the correlation between future values is 100%, but the correlation of exposures is only 96%. We can therefore see that the netting benefit depends not only on the correlation of future values, but also on the relative offset of the future values (in Table 13-2, trades 1 and 2 have positive and negative expected future values respectively which reduces the netting benefit). Netting not only depends on the *structural* correlation between the future values of trades, but also on the relative offset of those values from zero.

Consider the results shown in Table 13-4. Trade 1 has a strongly negative future value in all scenarios and therefore offsets the positive future value of trade 2 in scenarios 1–3. The EE is reduced from 9 to just 1. This is a result of trade 1 having an overall negative future value and is not solely linked to the structural correlation between the trades (indeed, the future values have been constructed to have zero correlation). For example, if the trade 1 future values are increased by +10 then the reduction is only 3, even though the correlation of future values is the same.

An illustration of the impact of negative future value of a netting set is shown in Figure 13-23. Negative future value will create netting benefit irrespective of the structural correlation between trades.

TABLE 13-4 Illustration of the Impact of Netting When There Is an Initial Negative Future Value. The expected exposure is shown, assuming each scenario has equal weight.

	Future Value		Total Exposure		
	Trade 1	Trade 2	No Netting	Netting	Netting Benefit
Scenario 1	−20	25	25	5	20
Scenario 2	−25	15	15	0	15
Scenario 3	−15	5	5	0	5
Scenario 4	−15	−5	0	0	0
Scenario 5	−25	−15	0	0	0
EE			9	1	

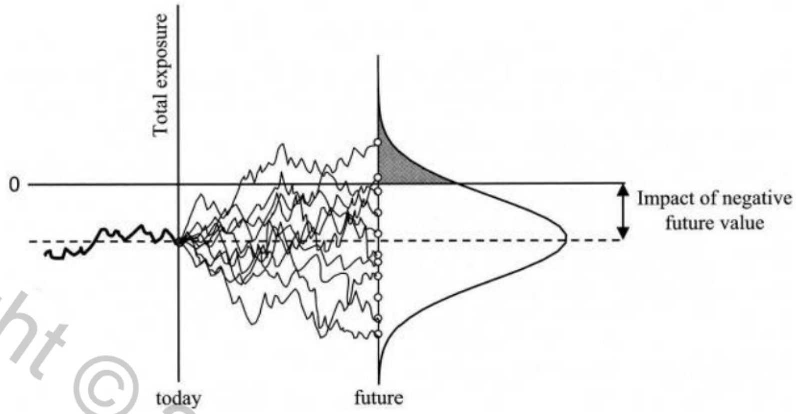


FIGURE 13-23 Schematic illustration of the impact of a negative future value.

TABLE 13-5 Illustration of the Impact of Netting When There Is a Positive Future Value. The expected exposure is shown, assuming each scenario has equal weight.

	Future Value		Total Exposure		
	Trade 1	**Trade 2**	**No Netting**	**Netting**	**Netting Benefit**
Scenario 1	45	15	60	60	0
Scenario 2	35	5	40	40	0
Scenario 3	25	−5	25	20	5
Scenario 4	15	−15	15	0	15
Scenario 5	5	−25	5	0	5
EE			29	24	5

A positive future value can be considered also to have a beneficial impact with respect to netting. Consider the results shown in Table 13-5. Trade 1 has positive future value in all scenarios, which nets with the negative future value of trade 2 in scenarios 4 and 5 even though the correlation of the trade 1 and trade 2 future values is 100%.

An illustration of the impact of the positive future value of a netting set is shown in Figure 13-24. It is important to emphasise that even highly correlated trades can give rise to netting benefits as the exposure may not be as highly correlated (as in the example in Table 13-2). A practical example of this could be two otherwise identical swaps but with different swap rates.

CREDIT EXPOSURE AND COLLATERAL

We now describe the key components in understanding the impact of collateral on credit exposure. Collateral typically reduces exposure but there are many (sometimes subtle) points that must be considered in order to assess properly the true extent of any risk reduction. To account properly for the real impact of collateral, parameters such as thresholds and minimum transfer amounts must be properly understood and represented appropriately. Furthermore, the "margin period of risk" must be carefully analysed to determine the true period of risk with respect to collateral transfer.

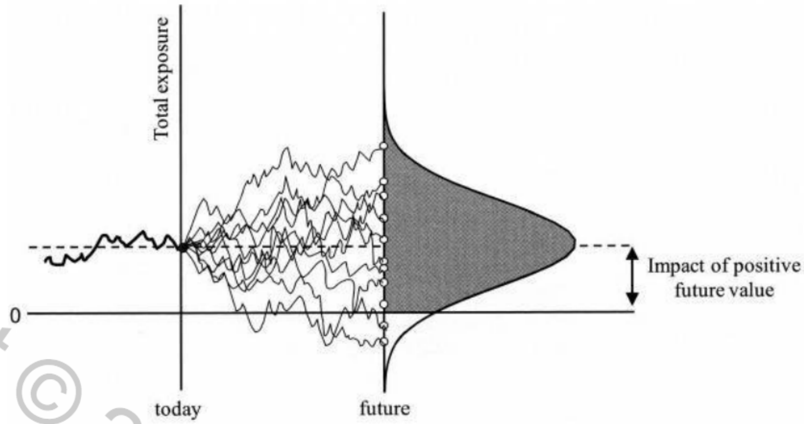


FIGURE 13-24 Schematic illustration of the impact of a positive future value.

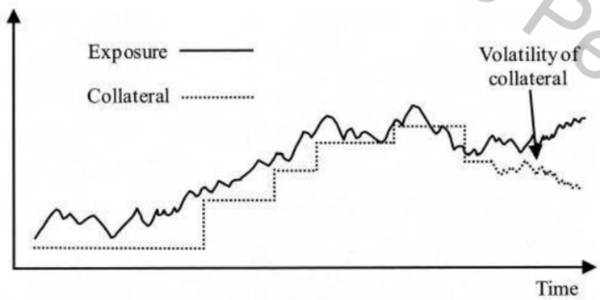


FIGURE 13-25 Illustration of the impact of collateral on credit exposure showing the delay in receiving collateral and the granularity receiving and posting collateral amounts discontinuously.

Also shown is the impact of the volatility of collateral itself (for ease of illustration this is shown in the last period only).

In addition to reducing it, collateral transforms counterparty risk into other risks, which must be thoroughly appreciated. Most notably, collateral leads to operational risk, legal risk and liquidity risk. Effective collateral management is counterproductive unless these risks are well understood and properly managed. We will highlight these risks in this section.

Collateralisation of credit exposure can substantially reduce counterparty risk but to quantify the extent of the risk mitigation is not trivial and requires many, sometimes subjective, assumptions. To the extent that collateral is not a perfect form of risk mitigation, there are three considerations, which are illustrated in Figure 13-25. Firstly, there is a granularity effect because it is not always possible to ask for all of the collateral required due to parameters such as thresholds and minimum transfer amounts (note that this can sometimes lead to a beneficial overcollateralisation as seen in Figure 13-25, where the collateral amount is for a short period greater than the exposure). Note that this must also consider the impact of collateral that an institution must themselves post. Secondly, there is a delay in receiving collateral which involves many aspects such as the operational components of requesting and receiving collateral to the possibility of collateral disputes. Thirdly, we must consider a potential variation in the value of the collateral itself (if it is not cash). We also emphasise that the treatment of collateral is path-dependent since the amount of collateral called for at a given time depends on the amount of collateral called (or posted) in the past. This is especially important in the case of two-way collateral agreements.

How Much Collateral?

The first question to ask is how much collateral may be requested at a given point in time. The parameters in a typical collateral support annex (CSA) do not, by design, aim for a continuous posting of collateral. This is because the operational cost and liquidity requirements of collateral posting are significant and one or both parties may find it beneficial to reduce such requirements within reason. The threshold and minimum transfer amount (discussed in Chapter 11) serve this purpose. The threshold is an amount below which collateral may not be called and the minimum transfer amount is the smallest amount that can be requested at a particular time. Note that in the case of a two-way CSA, both parties are subjected to the impact of thresholds and minimum transfer amounts.

The following steps define the amount of collateral required at a given time:

1. Add or subtract any specified independent amount to the market value of the trades (V).
2. Calculate the required collateral amount, taking into account the threshold using the formula

$$\max(V - \text{threshold}_c, 0) - \max(-V - \text{threshold}_I, 0) - C \quad \textbf{(13.6)}$$

   where V represents the current mark-to-market value[19] of the relevant trades, threshold_I and threshold_C represent the thresholds for the institution and their counterparty, respectively and C represents the amount of collateral held already. If the above calculation results in a positive value, then collateral can be requested, whilst a negative value indicates the requirement to post collateral (subject to the points below).
3. Determine whether the absolute value of the amount calculated above is above the minimum transfer amount. If not, then no call can be made.
4. If the amount is above the minimum transfer amount then round it to the relevant figure as specified in the CSA.

SPREADSHEET 13-6 Collateral Calculation

To download Spreadsheet 13-6, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 8 exercises.

Let us consider a collateral calculation assuming a two-way CSA with the threshold, minimum transfer amount and rounding equal to \$1,000,000, \$100,000 and \$25,000, respectively. Initially, we show an example in Table 13-6 where there is an exposure resulting in \$775,000 of collateral being called for. Whilst the mark-to-market of the underlying trades or "portfolio value" is \$1,754,858, the first million dollars of exposure cannot be collateralised due to the threshold. The required collateral is assumed to be rounded up to the final amount of \$775,000. Of course, assuming the counterparty agrees with all the calculations they will calculate a value of −\$775,000, meaning that they will post this amount.

TABLE 13-6 Example Collateral Calculation

	Collateral Calculation
Portfolio value	\$1,754,858
Collateral held	—
Required collateral (Equation (13.6))	\$754,858
Above minimum transfer amount?	YES
Rounded amount	\$775,000

TABLE 13-7 Example Collateral Calculation with Existing Collateral

	Collateral Calculation
Portfolio value	\$1,623,920
Collateral held	\$775,000
Required collateral (Equation (13.6))	−\$151,080
Above minimum transfer amount?	YES
Rounded amount	−\$150,000

In Table 13-7, the situation has changed since the collateral has been received and the exposure of the institution has dropped. The result of this is that they are required to post collateral back. Note that, whilst they still have uncollateralised exposure, they are required to do this because of the threshold, i.e., they must return collateral as their net exposure of \$848,920[20] has fallen below the threshold.

Margin Period of Risk

Now we consider how long it will take to receive collateral. This involves estimating the "margin period of risk," which is much more than the contractual time between collateral (margin) calls. Such a period is crucial since it

[19] In comparison to previously, this can be considered to be the mark-to-market value.

[20] This is the portfolio value of \$1,623,920 less the collateral held of \$775,000.

defines the length of time without receiving collateral where any increase in exposure will remain uncollateralised. It is important to model the exposure evolution over the margin period of risk to understand properly the impact of collateral. Where collateral is in a different currency and/or security, the variation in the FX rate and collateral price must also be accounted for, as this adds additional risk.

In order to assess the margin period of risk, it is important to consider all of the following effects that may slow down the collateral process:

- *Valuation/margin call.* This represents the time taken to compute current exposure and the current market value of collateral, working out if a valid call can be made and finally making that call. This should include the time delay due to the contractual period between margin calls (often daily calls are contractual but sometimes longer periods may apply).
- *Receiving collateral.* The delay between a counterparty receiving a collateral request (fax/email) to the point at which they release collateral. The possibility of a dispute (i.e., the collateral giver does not agree with the amount called for) should be incorporated here.
- *Settlement.* Collateral will not be received immediately, as there is a settlement period depending on the type of collateral. Cash collateral may settle on an intraday basis whereas other securities will take longer. For example, government and corporate bonds may be subject to 1-day and 3-day settlement periods, respectively.
- *Grace period.* In the event a valid collateral call is not followed by the receipt of the relevant collateral, there may be a relevant grace period before the counterparty would be deemed to be in default. This is sometimes known as the cure period.
- *Liquidation/closeout and re-hedge.* Finally, it will be necessary to liquidate collateral and close out and re-hedge positions.

We finally note that all of the above assessments should be considered in a scenario where the relevant counterparty is defaulting. This worst-case scenario is valid since one must base all calculations on the assumption that a counterparty *will* default, as discussed in Chapter 9. An institution is not concerned with the time taken to receive collateral in normal cases and normal market conditions (which may well be small) because collateral performs no function (in terms of mitigating counterparty risk at least[21]) in these situations. Instead, the institution must consider a scenario where their counterparty is in default and market conditions may be far from normal. In such a scenario, the time before being able to take delivery of collateral after a valid call (or alternatively to put the counterparty into default) can be significant.

Under Basel II rules, the minimum margin period of risk which must be assumed for OTC derivatives is 10 days (business), assuming collateral may be called for on a daily basis.[22] OTC derivatives and repo transactions are considered separately since they are governed by different documentation. Collateralisation in repo markets is generally tighter and the minimum period assumed is therefore lower due partly to the more complex nature of OTC derivatives, which makes valuation more complex. A possible scenario equating to such an assumption is shown in Table 13-8.

The above periods could easily be argued to be different depending on the precise assumptions and legal interpretations. Longer margin periods of risk could be appropriate depending on the collateral agreement and counterparty in question, as well as legal considerations and even the management structure of the institution concerned (institutions may be more lenient with certain counterparties to maintain good relations). In particular, Table 13-8 does not assess potential delays because of disputes or longer grace periods, which may be likely in practice. In particular, under Basel III, a margin period of risk of 20 days must be assumed in certain cases. An institution should decide carefully on the relevant margin period of risk with all of these considerations taken into account.

Clearly, the margin period of risk is a rather crude "catch-all" parameter. By definition, there is little empirical data on this parameter[23] and the correct modelling of collateral calls is complex. For example, in the case of a dispute, the protocol that should be followed is that the undisputed amount is transferred and then the parties involved enter into negotiations to agree on the disputed amount. The

[21] For example, in such a situation collateral may provide funding benefit.

[22] If this is not the case then the additional number of contractual days must be added to the time interval used.

[23] Although experiences such as the Lehman Brothers bankruptcy, where market participants typically took around 5–10 business days to close out portfolios, are reasonably consistent with this value.

TABLE 13-8 Example Timeline for the Margin Period of Risk in a Worst-case Scenario Based on the Assumption of a Daily Margin Call. This does not consider the additional delay potentially caused by disputes. The Basel II minimum period is also shown.

	OTC Derivatives (CSA[a])	Repo (GMRA[b])
Valuation/margin call	2 days	—
Receiving collateral	1 day	1 day
Settlement	2 days	1 day
Grace period	3 days	—
Liquidation/closeout and re-hedge	2 days	1 day
Total	10 days	3 days
Basel II minimum period	10 days	5 days

[a]Credit support annex.

[b]Global master repurchase agreement.

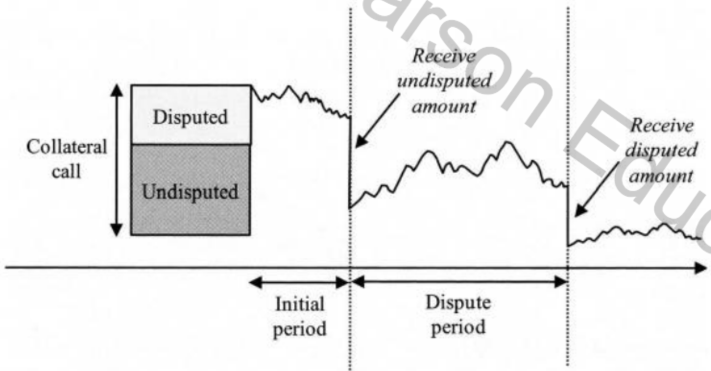


FIGURE 13-26 Illustration of the impact of a dispute on the margin period of risk assuming that the institution does eventually receive the full amount, including the disputed component.

We note that a fixed margin period of risk is a simple representation of the two periods above.

latter procedure may take some significant time, as experienced by many institutions during the financial crisis. This process is illustrated in Figure 13-26. In theory, receiving collateral should be divided into two parts, the undisputed and disputed amounts with associated periods. In practice, this is probably extraneous.

A simple example of the impact of collateral on exposure is given in Table 13-9 assuming a two-way CSA. In scenarios 1–3 the exposure is reduced significantly, since collateral is held. The exposure is not perfectly collateralised, which may be the case in practice due to factors such as thresholds and minimum transfer amounts. In scenario 4, the value of the portfolio is negative and collateral must therefore be posted, but this does not increase the exposure (again in practice due to aspects such as thresholds and minimum transfer amounts). Finally, in scenario 5, the posting of collateral *creates* exposure.[24] In comparison with the benefits shown in the other scenarios, this is not

[24] In practice, this can happen when previously posted collateral has not yet been returned as required.

TABLE 13-9 Illustration of the Impact of Collateral on Exposure. The expected exposure is shown assuming each scenario has equal weight.

	Future Value		Exposure		Benefit
	Portfolio	**Collateral**	**No Collateral**	**With Collateral**	
Scenario 1	25	23	25	2	23
Scenario 2	15	12	15	3	12
Scenario 3	5	3	5	2	3
Scenario 4	−5	−2	0	0	0
Scenario 5	−15	−18	0	3	−3

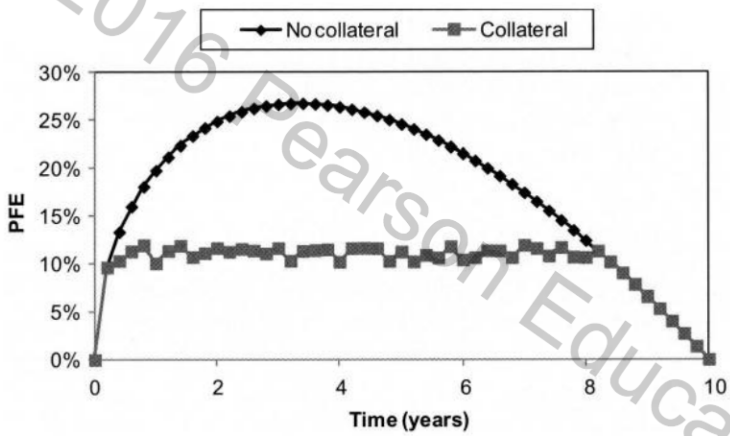


FIGURE 13-27 Illustration of the impact of collateral on exposure.

The collateral threshold is assumed to be 10%.

a particularly significant effect, but it is important to note that collateral can increase as well as reduce exposure.

Impact of Collateral on Exposure

Figure 13-27 shows an example of the impact of collateral on exposure. There are two main effects to notice. Firstly, the effect of a threshold is effectively to cap the exposure around the threshold amount. The collateral has little effect at the beginning and end of the profile where the exposure is relatively small. The second effect is the impact of the delay in receiving collateral; the need to post collateral and parameters such as minimum transfer amounts create some risk above the threshold.

Repos and Overcollateralisation

Repos represent collateralised loans. An institution borrowing cash and pledging some security as collateral is entering into a repo (repurchase agreement) whilst the other party is entering into a reverse repo. As a reverse repo is effectively a collateralised loan, it carries counterparty risk. The exposure on a reverse repo transaction is Exposure = Max(cash value − bond value, 0). To protect against counterparty risk, the collateral will attract a haircut meaning that the bond value at trade inception will exceed the amount of cash borrowed. In some repo transactions, variation margin (collateral) may be taken which will minimise the impact of changes in the bond price. Since a reverse repo is overcollateralised by the embedded haircut, the counterparty risk should be relatively small. There is a chance that a decline in the bond price (in-between variation margin calls) can create some counterparty risk. Figure 13-28 illustrates the exposure of a reverse repo transaction for different haircut levels assuming that variation margin is also used. The exposure (shown in terms of a bond notional of 100)

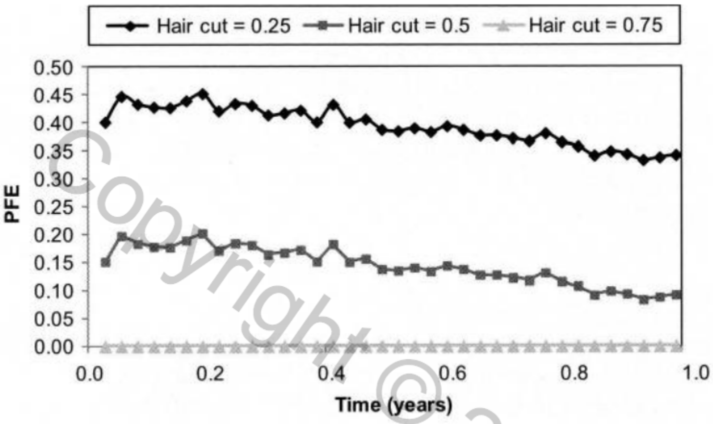


FIGURE 13-28 Illustration of the risk of a reverse repo transaction as a function of the haircut of maturity 1 year with the underlying collateral being a 5-year bond of notional amount 100.

The margin period of risk is assumed to be 5 days.

is quite small and almost negligible if the haircut is reasonably large.

The above example assumes a 1-year repo transaction. Whilst this is possible, repos are typically of much shorter term (e.g., overnight or 1 week). This and the over-collateralisation mean that the counterparty risk in reverse repos is typically small in comparison with other cases.

RISK-NEUTRAL OR REAL-WORLD?

A final consideration in terms of defining credit exposure is whether it should be done with respect to risk-neutral or real-world parameters. In the most simple terms, pricing (CVA) should use the former whilst risk management (PFE) the latter. However, the actual situation is more complicated.

The Importance of Measure

Scenario generation for risk management purposes and arbitrage pricing theory use different "measures." Arbitrage-based pricing uses the so-called *risk-neutral* measure, which is justified through hedging considerations. Parameters (and therefore probability distributions) such as drifts and volatilities are market-implied and need not correspond to the real distributions (or even comply with common sense). For a risk management application, one does not need to use the risk-neutral measure and should be focused rather on the *real-world measure*, estimated using, for example, historical data. Risk-neutral parameters are typically used in pricing applications (CVA), whilst real-world parameters generally form the basis of risk management models (PFE). This is the general distinction but there are necessary exceptions, which we discuss below.

The types of parameters to be considered are:

- *Drift*—the trend of market variables.
- *Volatility*—the future uncertainty of market variables.
- *Correlation*—the co-movement between market variables.

In addition to the above general definitions, effects like mean-reversion should be considered. Many market variables (for example, commodities and interest rates) tend to mean-revert over time, which pulls long-term rates back to some average level. Mean-reversion has an impact on future spot prices and volatilities. Risk-neutral mean-reversions, whilst often hard to calibrate, tend to be smaller than mean-reversions estimated from historical data.

Drift

A key difference between VaR analysis for market risk (for example) and credit exposure quantification for CVA purposes is the time horizon concerned. In the relatively short market risk horizon (for example, 10 days in VaR approaches), the drift of an underlying asset is of secondary importance vis-à-vis its volatility and is often ignored. However, in the longer time horizons required for assessing credit exposure and CVA, drift will be a key consideration alongside volatility. In other words, the *trend* of an underlying variable can be just as important as its *uncertainty*.

One area where risk-neutral parameters tend to be used even for risk management simulations is the determination of the drifts of underlying risk factors, which are typically calibrated from forward rates. The consideration of drifts is important since the impact of volatility approximately follows the square root of time scaling whereas the drift scales more linearly—so in the end a strong drift will eventually dominate. Futures (or equivalently forward) prices have long been an important mechanism of price discovery in financial markets as they represent the intersection of expected supply and demand at some future point in time. Forward rates can sometimes be very far from spot rates and it is important to understand whether or not this is truly the "view of the market." Some important technical factors are:

- *Commodity prices.* In addition to market participants' view of the direction of commodity prices, storage costs (or lack of storage), inventory and seasonal effects can move commodities futures apart from spot rates. For high inventories the futures price is higher than the spot price (contango). When inventories are low, commodity spot prices can be higher than futures prices (backwardation).
- *Interest rates.* Yield curves may be upwards-sloping or downwards-sloping (and a variety of other shapes) due to the risk appetite for short-, medium- and long-term interest rate risk and the view that rates may increase or decrease.
- *Credit spreads.* Credit curves may be increasing or decreasing either due to demand for credit risk at certain maturities or the view that default probability will be increasing or decreasing over time.
- *Foreign exchange (FX) rates.* Forward FX rates are determined from an arbitrage relationship between the interest rate curves for the relevant currency pair. Expectation of future FX rates may have an influence on the current interest rate curves in the corresponding currencies. For example, FX forward rates are determined by a differential in the underlying interest rates. There has long been doubt regarding the ability of long-term forward rates to predict future spot rates; see, for example, Meese and Rogoff (1983) and a review by Sarno and Taylor (2002).

There has been much empirical testing of the relationship between spot and futures prices across different markets. It is a generally held belief that the futures price is a biased forecast of the future spot price, contrary to the efficient market hypothesis. If we take the view that the forward rate is the best expectation of the future spot rate then this may lead to a strong drift assumption. If this assumption is wrong then it will significantly overstate or understate the risk.

Despite the above problems with drifts, most PFE and CVA calculations will calibrate to forward rates in the market. From the CVA point of view, this is justified by hedging and more discussion is given. For PFE purposes, this is often done more for convenience's sake, since it means that simple instruments are by construction priced properly and circumvents the need to attempt to estimate the "real-world" drift of risk factors.

The key point to take away is that markets are imperfect and so we cannot always expect current futures prices to be the best estimate of spot prices in the future. We should bear this in mind when assessing and pricing counterparty risk, especially for long time horizons. Advocating the estimation of real-world drifts is not the intention here. However, it is important to be aware of the implications of using risk-neutral drifts for PFE quantification and for CVA calculations when hedging is not perfect.

Example

Consider a transaction whose future value has a volatility of 10% and a drift of 5% over 1 year.

The expected exposure based on the usual formula is

$$[5\% \times \Phi(5\%/10\%) + 10\% \times \varphi(5\%/10\%)] = 6.98\%$$

On the other hand, consider the reverse transaction. The expected drift would be -5% and the expected exposure

$$[-5\% = \Phi(-5\%/10\%) + 10\% \times \varphi(-5\%/10\%)] = 1.98\%$$

Is it correct that the first transaction has a CVA that is approximately three and a half times greater than the second?

Volatility

To quantify exposure, one might use a historical estimate of volatility. However, to calculate CVA, implied volatilities are more relevant. Again, there is the caveat related to the extent to which the volatility component of CVA can (and will) be hedged. We also note that (positive) mean-reversion has the effect of reducing long-term volatilities and thus is an important parameter to estimate.

If one uses a historical estimate of volatility, then the implicit assumption is that the past will be a good indication of the future. It is also necessary to decide what history of data to use; a short history will give poor statistics whereas a long history will give weight to "old" meaningless data. In quiet markets, the lack of volatility in historical time series will give low-risk numbers which may be misleading (recent changes to Basel capital rules require always using a stress period of data to overcome this), creating procyclicality. When markets suddenly become more volatile, the historical estimate will only gradually increase to reflect this as the window of data moves.

For most markets, there is likely to be implied volatility information, potentially as a function of strike and the maturity of the option. Implied volatility which will react quickly when the market becomes more uncertain and may be justified via the "market knows best" (or at least the market knows better than historical data). However, risk premiums embedded in market-implied volatilities will lead to a systematic overestimate of the overall risk. It has been argued that implied volatility is a superior estimator of future volatility (e.g., see Jorion, 2007, chapter 9) compared with historical estimation via time series approaches. The stability of the volatility risk premium and the fact that an overestimate of volatility will always lead to a more conservative[25] risk number give greater credence to this idea.

Correlations

Whilst it is at least conservative to assume volatilities are high, the same is not true of other quantities. When estimating correlation for modelling exposure, there may not be an obvious way of knowing whether a high or low (or positive or negative) value is more conservative. Indeed, in a complex portfolio it may even be that the behaviour of the exposure with respect to correlation is not monotonic.[26] Therefore, the use of some market-implied parameters cannot be justified on the basis that the resulting risk numbers will be conservatively high.

Implied correlations are sometimes available in the market. For example, a quanto option has a payoff in a different currency and thus gives information on the implied correlation between the relevant FX rate and the underlying asset. One key aspect of correlation is to determine wrong-way risk. For example, a quanto CDS (a CDS where the premium and default legs are in different currencies) potentially gives information on the correlation between the relevant FX rate and the credit quality of the reference entity in the CDS.[27]

Whilst implied correlation can sometimes be calculated, for most quantities no market prices will be available.

[25] Using implied volatility might be expected to produce an upwards bias due to a risk premium, leading to higher (more conservative) risk numbers.

[26] Meaning, for example, that the worst correlation may not be +100% or −100% but somewhere in-between.

[27] Assuming we can also observe the premiums of the "vanilla" or single-currency CDS.

This also means that the sensitivity of CVA to correlation parameters cannot generally be hedged and historical data will probably be used. A sensitivity analysis of correlation will be useful to understand the importance of a particular correlation parameter.

Conclusion

In summary, exposure quantification for risk management (PFE) should generally focus on real parameters, with market-implied parameters used when there are good reasons (such as in the example of using drifts and implied volatility above). Exposure quantification for pricing (CVA) should *generally* focus on (risk-neutral) market-implied parameters. An obvious exception here would be the need to use historical correlations since market-implied parameters are typically not observed. The potential difference between real-world and risk-neutral exposure will be discussesd in Chapter 15.

SUMMARY

In this chapter we have discussed credit exposure. Some key definitions of potential future exposure, expected exposure and expected positive exposure have been given. The factors impacting future exposures have been explained and we have discussed the impact of netting and collateral.

APPENDIX A

Formulas for EE, PFE and EPE for a Normal Distribution

Consider a normal distribution with mean μ (expected future value) and standard deviation (of the future value) σ. Let us calculate analytically the two different exposure metrics discussed. Under the normal distribution assumption, the future value of the portfolio in question (for an arbitrary time horizon) is given by:

$$V = \mu + \sigma Z,$$

where Z is a standard normal variable.

i) Potential future exposure (PFE)

This measure is exactly the same as that used for value-at-risk calculations. The PFE at a given confidence level α, (PFE_α) tells us an exposure that will be exceeded with a probability of no more than $1 - \alpha$. For a normal distribution, it is defined by a point a certain number of standard deviations away from the mean:

$$PFE_{\alpha} = \mu + \sigma\Phi^{-1}(\alpha),$$

where $\Phi^{-1}(.)$ represents the inverse of a cumulative normal distribution function (this is the function NORMSINV(.) in Microsoft Excel™). For example, with a confidence level of $\alpha = 99\%$, we have $\Phi^{-1}(99\%) = +2.33$ and the worst case exposure is 2.33 standard deviations above the expected future value.

ii) Expected exposure (EE)

Exposure is given by:

$$E = \max(V,0) = \max(\mu + \sigma Z,0)$$

The EE defines the expected value over the positive future values and is therefore:

$$EE = \int_{-\mu/\sigma}^{\infty} (\mu + \sigma x)\varphi(x)dx = \mu\Phi(\mu / \sigma) + \sigma\varphi(\mu / \sigma),$$

where $\varphi(.)$ represents a normal distribution function and $\Phi(.)$ represents the cumulative normal distribution function. We see that EE depends on both the mean and the standard deviation; as the standard deviation increases so will the EE. In the special case of $\mu = 0$ we have $EE_0 = \sigma\varphi(0) = \sigma/\sqrt{2\pi}$ 0.40σ.

iii) Expected positive exposure

The above analysis is valid only for a single point in time. Suppose we are looking at the whole profile of exposure defined by $V(t) = \sigma\sqrt{t}Z$ where σ now represents an annual standard deviation (volatility). The EPE, integrating over time and dividing by the time horizon, would be:

$$EPE = \frac{1}{\sqrt{2\pi}}\sigma\int_0^T \sqrt{t}dt / T = \frac{2}{3\sqrt{2\pi}}\sigma T^{1/2} = 0.27\sigma T^{1/2}.$$

APPENDIX B

Simple Netting Calculation

We have already shown in Appendix A that the EE of a normally distributed random variable is:

$$EE_i = \mu_i\Phi(\mu_i/\sigma_i) + \sigma_i\varphi(\mu_i/\sigma_i).$$

Consider a series of independent normal variables representing transactions within a netting set (NS). They will have a mean and standard deviation given by:

$$\mu_{NS} = \sum_{i-1}^{n}\mu_i \qquad \sigma_{NS}^2 = \sum_{i=1}^{n}\sigma_i^2 + 2\sum_{\substack{i=1\\ j>1}}^{n}\rho_{ij}\sigma_i\sigma_j$$

where ρ_{ij} is the correlation between the future values. Assuming normal variables with zero mean and equal standard deviations, $\bar{\sigma}$, we have that the overall mean and standard deviation are given by:

$$\mu_{NS} = 0 \qquad \sigma_{NS}^2 = \left(n + n(n-1)\bar{\rho}\right)\bar{\sigma}^2,$$

where $\bar{\rho}$ is an average correlation value. Hence, since $\varphi(0) = 1/\sqrt{2\pi}$, the overall EE will be:

$$EE_{NS} = \bar{\sigma}\sqrt{n + n(n-1)\bar{\rho}} / \sqrt{2\pi}$$

The sum of the individual EEs gives the result in the case of no netting (NN):

$$EE_{NN} = \bar{\sigma}n / \sqrt{2\pi}$$

Hence the netting benefit will be:

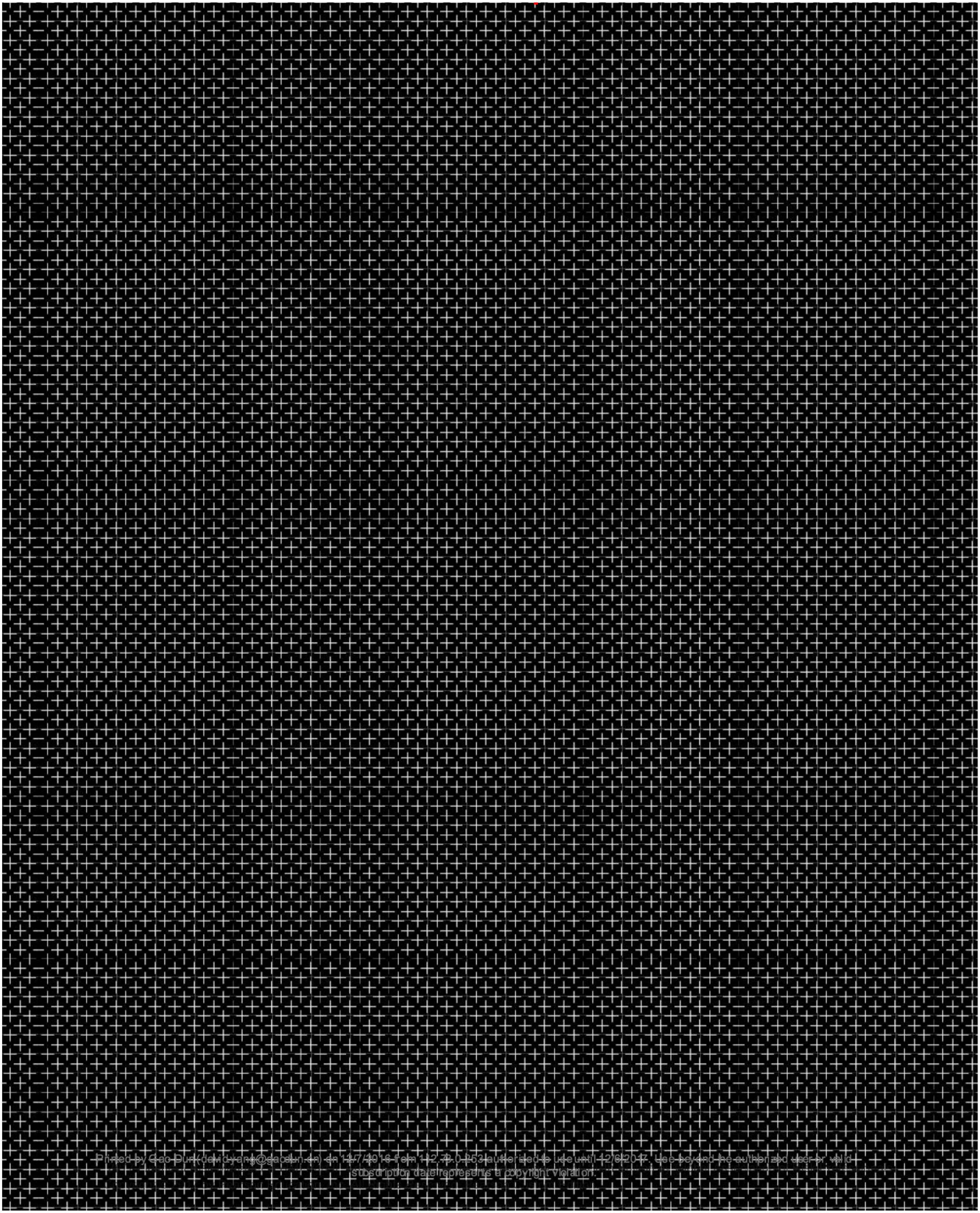
$$EE_{NS} / EE_{NN} = \frac{\sqrt{n + n(n-1)\bar{\rho}}}{n}$$

In the case of perfect positive correlation, $\bar{\rho} = 100\%$, we have:

$$EE_{NS} / EE_{NN} = \frac{\sqrt{n + n(n-1)}}{n} = 100\%$$

The maximum negative correlation is bounded by $\bar{\rho} \geq -1/(n-1)$ and we therefore obtain:

$$EE_{NS} / EE_{NN} = \frac{\sqrt{n - n(n-1)/(n-1)}}{n} = 0\%$$



Default Probability, Credit Spreads and Credit Derivatives

14

Learning Objectives

After completing this reading you should be able to:

- Distinguish between cumulative and marginal default probabilities.
- Calculate risk-neutral default probabilities, and compare the use of risk-neutral and real-world default probabilities in pricing derivative contracts.
- Compare the various approaches for estimating price: historical data approach, equity based approach, and risk neutral approach.
- Describe how recovery rates may be estimated.
- Describe credit default swaps (CDS) and their general underlying mechanics.
- Describe the credit spread curve and explain the motivation for curve mapping.
- Describe types of portfolio credit derivatives.
- Describe index tranches, super senior risk, and collateralized debt obligations (CDO).

Excerpt is Chapter 10 of Counterparty Credit Risk and Credit Value Adjustment, *Second Edition, by Jon Gregory.*

Creditors have better memories than debtors.

—Benjamin Franklin (1706–1790)

So far, this book has been largely concerned with credit exposure, the market risk component of counterparty risk. Now we focus on the credit risk component arising from the probability of counterparty default and the loss incurred as a result. We will also discuss recovery rates, which define the amount of a claim that is received from a defaulted counterparty.

Default probability plays a critical role in counterparty risk assessment and valuation (CVA). There are different ways to define default probability, which we will explain, noting the important difference between using real-world (e.g., historical data) and risk-neutral probabilities (market data). In the latter case, we consider mapping methods that may be used for estimating the credit spread of a counterparty where this cannot be estimated directly. We also consider the term structure of default probability (how default probability changes over time) and show that this is an important consideration. The empirical relationship between real-world and risk-neutral default probabilities (a very important point for defining CVA) is discussed.

Finally, we will examine single-name and portfolio credit derivative products, which are important for hedging purposes and the consideration of wrong-way risk (Chapter 16).

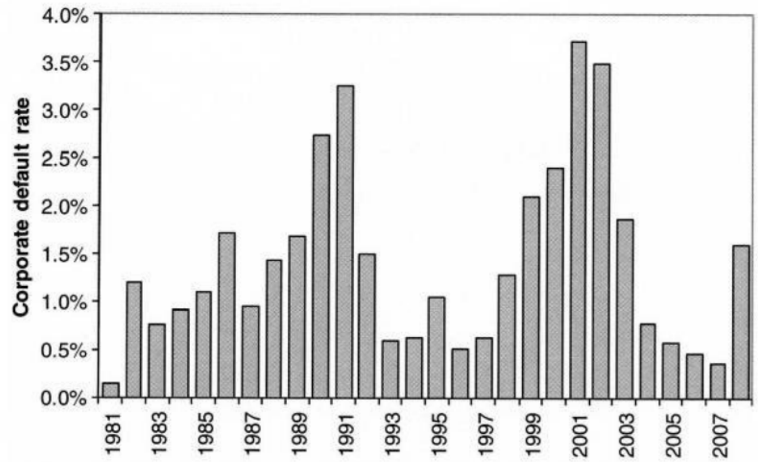


FIGURE 14-1 Corporate annual default rates (average of investment and speculative grade rates).

Source: Standard & Poor's (2008).

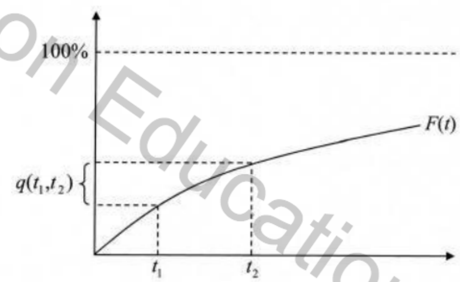


FIGURE 14-2 Illustration of cumulative default probability function, $F(t)$, and marginal default probability, $q(t_1, t_2)$.

DEFAULT PROBABILITY AND RECOVERY RATES

An example of historical default rates for investment- and speculative-grade assets is shown in Figure 14-1, illustrating that default rates tend to vary substantially through the economic cycle.

Defining Default Probability

In Appendix 14A we define default probability in more mathematical detail. We refer to the cumulative default probability, $F(t)$, which gives the probability of default any time from now (assuming the counterparty is not currently in default) until time t. This is illustrated in Figure 14-2. The function must clearly start from zero and tend towards 100% (every counterparty defaults eventually!). A *marginal* default probability, which is then the probability of a default between two specified future dates, is given by

$$q(t_1, t_2) = F(t_2) - F(t_1) \qquad (t_1 \le t_2). \tag{14.1}$$

We can see that $F(.)$ must be monotonically increasing to avoid marginal default probabilities being negative.

Real and Risk-Neutral Default Probabilities

It is well known in finance that there is a difference between a real-world parameter (for example, the historical volatility of an asset) and a risk-neutral one (for example, the implied volatility derived from an option price). Real-world (also known as physical) parameters

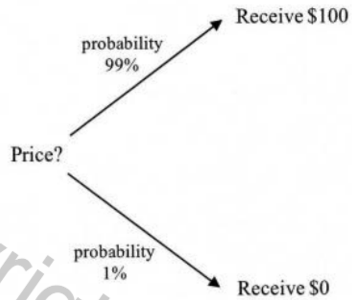


FIGURE 14-3 Illustration of the difference between real and risk-neutral default probabilities.

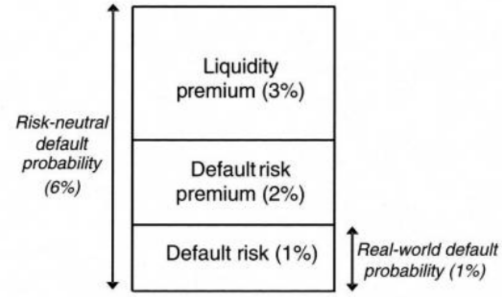


FIGURE 14-4 Example illustration of different components of a bond price and the difference between real and risk-neutral default probabilities.

aim to reflect the true value of some financial underlying whilst risk-neutral parameters reflect parameters derived from market prices. The distinction between real and risk-neutral parameters is important.

For our current purpose, real default probabilities will be the assessment of future default probability for the purposes of risk management or other analysis. Risk-neutral default probabilities are not estimates of the actual probabilities of default but rather reflect the market price of default risk.

There is reason to expect that real and risk-neutral default probabilities will be very different. This can be understood from a simple example illustrated in Figure 14-3, which is a bet in which there are two possible outcomes: a gain of $100 with 99% probability or a zero gain with 1% probability. This example is equivalent to a zero-recovery, zero-coupon bond in a zero-interest-rate environment. A quick calculation would suggest that the price of the bond is $99 (99% × $100 + 1% × $0). However, no rational investor would enter into such a bet for $99 as the expected payoff is no greater and there is the risk of losing the entire stake.

Rational investors are risk-averse and would never accept risk without the expectation of making a positive return. Suppose an investor was willing to pay only $97 for the "bond" in the example above. They are (quite rationally) expecting a $2 reduction as compensation for the uncertainty of the return. We could call this a *default risk premium,* i.e., the premium that investors need in order to accept default risk. This would require probabilities in Figure 14-3 of 97% and 3%. Furthermore, suppose the investor is worried about the liquidity of the bond above as they may need or want to sell it at some point in the future. For this reason, they may only pay $94 for the bond. The further $3 could be described as a *liquidity premium.* The probabilities would now have to be 94% and 6%. These are not the real default probabilities, but rather constructed risk-neutral probabilities to make the numbers balance assuming that investors have no aversion to risk and will therefore take on the fair bet that Figure 14-3 will then represent. If $94 were the market price of the bond then the risk-neutral default probability would be 6%. We emphasise that this is an artificial probability derived from the market price and has nothing to do with the actual likelihood of the bond defaulting (which is 1%), as illustrated in Figure 14-4.

It is important to understand that a difference in real-world and risk-neutral default probabilities is not conflicting and simply represents a difference in what they represent. Indeed, in line with the above, Altman (1989) tracks the performance of portfolios of corporate bonds for a given rating and finds that the returns outperform a risk-free benchmark (which is a portfolio of Treasury bonds). The reason for the outperformance is that the return on the corporate bonds is more than adequate to cover the default losses experienced. This shows clearly that bond investors are being compensated for components above expected default losses and that the size of these components is significant. Risk-neutral default probabilities are materially higher than real-world ones.

There is no conflict between risk-neutral and real default probabilities. Real-world default probabilities are the actual assessment of the probability of a counterparty defaulting, which is therefore relevant for any quantitative

TABLE 14-1 1-Year Transition Matrix for Moody's Ratings

	AAA	AA	A	BBB	BB	B	CCC	Default
AAA	91.61%	7.70%	0.66%	0.00%	0.02%	0.00%	0.00%	0.01%
AA	1.13%	91.29%	7.21%	0.27%	0.06%	0.02%	0.00%	0.02%
A	0.07%	2.84%	91.30%	5.14%	0.51%	0.09%	0.02%	0.03%
BBB	0.05%	0.20%	5.15%	88.83%	4.54%	0.81%	0.24%	0.18%
BB	0.01%	0.06%	0.42%	6.25%	82.95%	8.48%	0.63%	1.20%
B	0.01%	0.05%	0.18%	0.39%	6.21%	81.93%	6.23%	5.00%
CCC	0.00%	0.03%	0.03%	0.19%	0.73%	11.22%	68.57%	19.23%

Source: From Tennant et. al. (2008)

assessment of return or risk management approach. Risk-neutral default probabilities reflect the market price and are therefore relevant for hedging purposes. Let us discuss the methods for estimating both and then return to the question of defining this difference, which will be discussed later.

We note that the above discussion applies to risk-neutral default probabilities from bond prices, but similar behaviour should be expected with respect to CDS-implied default probabilities. This is discussed in more detail later.

Estimating Real Default Probabilities—Historical Data

The most obvious assessment of real default probability comes from examining historical data and using past default experience to predict future default likelihood. For example, in Table 14-1 we show a transition matrix based on many years of data as published in Tennant et. al. (2008). This matrix gives the historical probability of moving from a given rating (on the left-hand column) to another rating (in the top row) during a period of one year. It also defines the default probabilities in the far right column. For example, the probability of an A rating being downgraded to BBB after one year is 5.14% and the chance of it defaulting is 0.03%.

Not only does Table 14-1 give information on the probability of default, it also provides greater structure for defining how defaults occur. For example, we see that a Triple-A credit has only a 0.01% chance of defaulting in a year but a 7.7% chance of deteriorating to a Double-A credit. A Triple-C has a large 19.23% chance of default but a 12.2% chance[1] of improving in credit rating over a year.

By making several assumptions, we can derive the cumulative default probabilities, $F(.)$, for each credit rating from Table 14-1. The main assumption[2] required in order to do this is that the matrix is constant through time. This is clearly a naïve assumption, as default and transition probabilities would be expected to change through the economic cycle, but it is reasonable for estimating default probabilities over long periods. Under such assumptions, we can simply multiply this matrix by itself $n-1$ times to derive an n-year matrix. The resulting cumulative default probabilities are shown in Table 14-2 and plotted in Figure 14-5.

SPREADSHEET 14-1

Historical Default Probabilities

To download Spreadsheet 14-1, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 10 exercises.

Looking at the results, apart from the obvious conclusion that firms with good credit ratings default less often

[1] The sum of the first six numbers in the bottom row of Table 14-1, which represent the total probability of an upgrade.

[2] Other assumptions are that in the data, only a maximum of one credit rating move was experienced in a given year and that credit ratings have no "memory"—e.g., a given rating that has been upgraded or downgraded recently is not different from the same rating not subject to such a move.

TABLE 14-2 Cumulative Default Probabilities Implied from the 1-Year Transition Matrix Shown in Table 14-1

	AAA	AA	A	BBB	BB	B	CCC
1	0.01%	0.02%	0.03%	0.18%	1.20%	5.00%	19.23%
2	0.02%	0.04%	0.08%	0.48%	2.75%	10.37%	32.99%
3	0.03%	0.07%	0.16%	0.90%	4.60%	15.72%	43.03%
4	0.05%	0.10%	0.27%	1.43%	6.68%	20.85%	50.54%
5	0.07%	0.15%	0.42%	2.06%	8.92%	25.65%	56.27%
6	0.09%	0.20%	0.60%	2.78%	11.26%	30.09%	60.77%
7	0.11%	0.27%	0.82%	3.58%	13.65%	34.15%	64.36%
8	0.14%	0.35%	1.09%	4.45%	16.05%	37.85%	67.30%
9	0.18%	0.45%	1.39%	5.39%	18.43%	41.22%	69.75%
10	0.22%	0.57%	1.73%	6.38%	20.76%	44.28%	71.83%

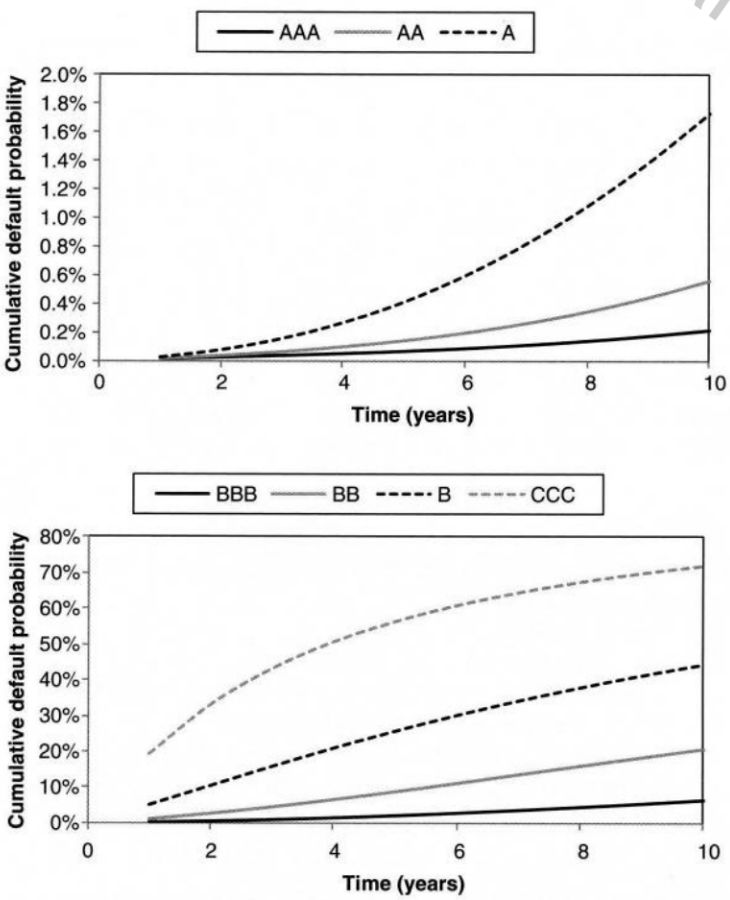


FIGURE 14-5 Illustration of cumulative default probabilities, as given in Table 14-2.

than those with worse ratings, we can also notice the following.

- *Investment grade credits.* These tend to have default probabilities that increase over time. For example, the 5-year Single-A (A) default probability is 0.42% but the 10-year probability is 1.73%, which is more than four times bigger.
- *Non-investment grade credits.* These credits tend to show the reverse effect, with default probabilities that increase much less strongly over time. For example, the 2-year Triple-C default probability is less than double the 1-year one (32.99% compared with 19.23%[3]).

The results can be explained by the mean-reversion of credit ratings, where above-average counterparties tend to deteriorate and vice versa. Hence (conditioning on no prior default), a good credit quality counterparty is much more likely to be downgraded than upgraded, whilst the reverse is true for a counterparty with a low rating. Such

[3] A key point to consider is that poor credit quality firms have default probabilities concentrated in the short term, not necessarily because their credit quality is expected to improve over time but instead because default in a future period can only be achieved by surviving to the start of this period. However, all other things being equal, we would expect the 2-year default probability to be 19.23% + 19.23% × (1 − 19.23%) = 34.76%. The actual 2-year default probability is less than this, illustrating that there is still a reduction in default probability and there is another component to be considered.

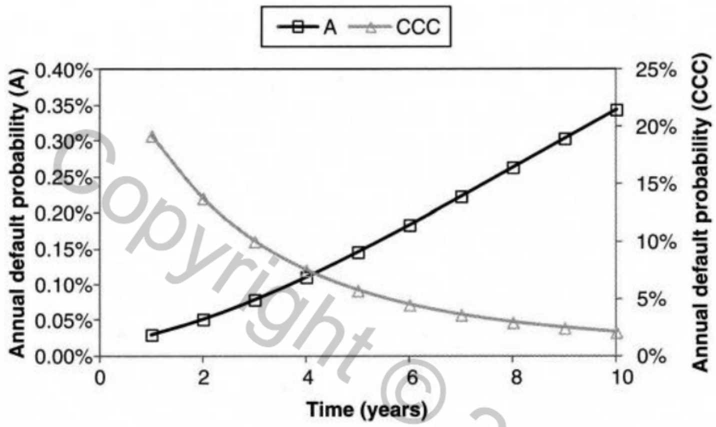


FIGURE 14-6 Annual historical default probabilities for Moody's A-rated firms computed using the data in Table 14-2.

For example, the point at 3 years represents the default probability in the interval 2 years to 3 years, which is 0.16% − 0.08% = 0.08%.

trends can easily be seen when looking at transition matrices as shown in Table 14-1. For example, the probabilities of being upgraded and downgraded from A are respectively 2.91% and 5.76%, whilst the equivalent numbers for CCC are 12.2% and zero.

In computing CVA, not only will the cumulative default probability be important, but also so will the way in which this is distributed marginally. We illustrate this in Figure 14-6, which shows annual default probabilities for A- and CCC-rated credits. The former increases significantly through time and the latter reduces. If an A-rated credit defaults then it is likely to be towards the end of the 10-year horizon considered, whilst a CCC-rated credit is likely to default much earlier.

Estimating Real Default Probabilities—Equity-Based Approaches

Equity-based approaches aim to estimate default probability from stock market information. In the classic Merton (1974) framework, the value of a firm (asset value) is considered stochastic and default is modelled as the point where the firm is unable to pay its outstanding liabilities when they mature. The original Merton model assumes that a firm has issued only a zero-coupon bond and will not therefore default prior to the maturity of this debt. Through option-pricing arguments, Merton then provides a link between corporate debt and equity via pricing formulas based on the value of the firm and its volatility (analogously to options being valued from spot prices and volatility). The problem of modelling default is transformed into that of assessing the future distribution of firm value and the barrier where default would occur. Such quantities can be estimated non-trivially from equity data and capital structure information. A key contribution of the Merton approach is that low-frequency binary events can be modelled via a continuous process and calibrated using high-frequency equity data.

KMV™ (now Moody's KMV) developed the Merton-style approach (e.g., see Kealhofer and Kurbat, 2002; Kealhofer, 2003) with the aim of predicting default via the assessment of 1-year default probability defined as EDF™ (expected default frequency). The KMV approach relaxed many of the stylised Merton assumptions. Their approach can be summarised broadly in three stages:

- estimation of the market value and volatility of a firm's assets;
- calculation of the distance to default, which is a standardised measure of default risk;
- scaling of the distance to default to the actual probability of default using a historical default database.

The distance to default (DD) measure is a standardised distance a firm is away from default. A key element of the KMV approach is to recognise the model risk inherent in this approach and rather to estimate the default probability empirically from many years of default history (and the calculated DD variables). For a firm with a DD of 4.0 (say), the question KMV attempt to answer is how often firms with the same DD have defaulted historically. The answer is likely to be considerably higher than the theoretical result of 0.003%.[4] This mapping of DD to actual default probability could be thought of as an empirical correction for model error. Note that, although the KMV approach relies on historical data, the EDF measure will still be dynamic due to constantly changing equity data.

A more recent and related, although simpler, approach is CreditGrades™. The aims of CreditGrades are rather similar to those of KMV, except that the modelling framework (see Finger et. al., 2002) is rather simpler and more transparent; in particular, there is no use of empirical data in order to map to an eventual default probability. In

[4] This arises since $\Phi^{-1}(-4.0) = 0.003\%$.

CreditGrades, default probability is defined by a simple formula with just a few model parameters.

Equity-based models for default probabilities have a place due to their ability to define default probability dynamically. This can be an advantage in a situation where historical default probabilities are considered too static a measure whilst probabilities defined directly from the credit market (discussed next) may be considered highly volatile and conservative due to the embedded default risk and liquidity premiums.

Estimating Risk-Neutral Default Probabilities

Risk-neutral default probabilities are those derived from credit spreads observed in the market. There is no single definition of a credit spread and it may be defined in slightly different ways and with respect to different rates. Common ways to define a credit spread are:

- from the premiums of single-name CDSs;
- from the traded spreads of asset swaps;[5]
- from bond prices, typically compared with some benchmark such as a treasury or LIBOR curve;
- using some proxy or mapping method.

All of the above are (broadly speaking) defining the same quantity, but small differences do exist in practice. The basis between CDS and bond spreads will be discussed later. For now, we will focus on deriving risk-neutral default probabilities from the defined credit spread, which reflects the market price of credit risk.

Suppose a counterparty has a certain constant probability of default each year of (say) 10%. This must be the conditional default probability (i.e., the default probability assuming default has not yet occurred) as otherwise, after more than 10 years, the total default probability would be greater than 100%. This is illustrated in Figure 14-7. The probability of defaulting in the second year is equal to the probability of surviving the first year and defaulting the next, which would be 90% × 10% = 9%. The probability of defaulting at any time in the first two years is then 10% + 9% = 19%. By similar arguments, the probability of defaulting in the third year must be 90% × 90% × 10% = 8.1%, and so on.

[5] An asset swap is essentially a synthetic bond, typically with a floating coupon.

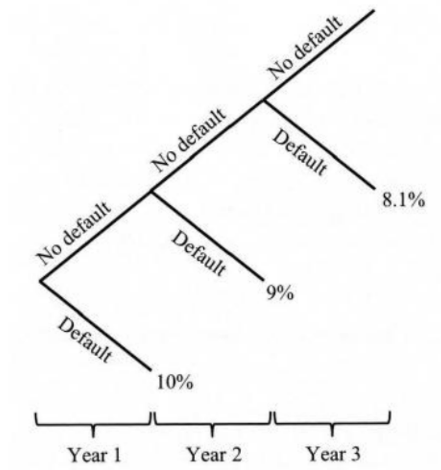


FIGURE 14-7 Illustration of the default process through time assuming a conditional default probability of 10% per year.

A more formal mathematical description is that default is driven by a Poisson process and the default probability for a future period u is given by

$$F(u) = 1 - \exp(-hu), \tag{14.2}$$

where h defines the hazard rate of default, which means that the conditional default probability in a period of length dt is given by $h * dt$. By choosing a hazard rate of 10.54%,[6] we can reproduce the results corresponding to the 10% annual default probability; for example, $1 - \exp(-10.54\% \times 2) = 19\%$ is the default probability in the first two years.

An approximate[7] relationship between the hazard rate and credit spread is

$$h \approx \frac{\text{spread}}{(1-\text{recovery})}, \tag{14.3}$$

where the assumed recovery rate is a percentage. Combining the above two equations gives the following approximate expression for risk-neutral default probability up to a given time u:

$$F(u) = 1 - \exp\left[-\frac{\text{spread}}{(1-\text{recovery})}u\right]. \tag{14.4}$$

[6] This can be found from inverting Equation (14.2) at the 1-year point as −log(1 − 10%), where log represents the natural logarithm.

[7] This assumes that the credit spread term structure is flat (credit spreads for all maturities are equal) and that CDS premiums are paid continuously, as discussed in Appendix B.

TABLE 14-3 Comparison of Simple and Accurate Calculations for an Example Credit Curve

Tenor	Spread	Hazard Rate (Approximation from Equation 14.3)	F(u)	Annual Default Probability (Equation (14.5))	Annual Default Probability (Exact)
1Y	300 bps	5.00%	4.88%	4.88%	4.82%
2Y	350 bps	5.83%	11.01%	6.13%	6.14%
3Y	400 bps	6.67%	18.13%	7.12%	7.28%
4Y	450 bps	7.50%	25.92%	7.79%	8.22%
5Y	500 bps	8.33%	34.08%	8.16%	8.93%

The reason that risk-neutral default probability depends on recovery can be explained as follows. Suppose a bond will default with a probability of 2% but the recovery value would be 50%. The expected loss is 1%, which is the same as if the bond had a 1% probability of default but the recovery value was zero. In the market we see only a single parameter (the credit spread) and must imply two values from it. Common practice is then to fix the recovery rate and derive the default probability. A higher recovery must be balanced (good for the bondholder) by a larger assumed default probability (bad for the bondholder).

The above formula is a good approximation generally, although to compute the implied default probabilities accurately we must solve numerically for the correct hazard rate, assuming a certain underlying functional form. The reader is referred to Appendix B and O'Kane (2008) for a more detailed discussion. Such an approach is also required to take into account the term structure of credit spreads and incorporate other aspects such as the convention of using upfront premiums in the CDS market.

SPREADSHEET 14-2

Implied Default Probabilities

To download Spreadsheet 14-2, visit http://www.cvacentral.com/books/credit-value-adjustment/spreadsheets and click Chapter 10 exercises.

In terms of marginal default probability between dates t_{i-1} and t_i, an obvious approximation would be to take the difference between the relevant cumulative default probabilities in Equation (14.4), leading to

$$q(t_{i-1}, t_i) \approx \exp\left[-\frac{\text{spread}_{t_{i-1}}}{(1-\text{recovery})}t_{i-1}\right] - \exp\left[-\frac{\text{spread}_{t_i}}{(1-\text{recovery})}t_i\right]. \quad (14.5)$$

This approach is used to define CVA under Basel III. It is only an approximation because it does not account for the shape of the credit curve prior to the time t_{i-1} (and the more sloped the curve is, the worse the approximation). We compared the simple formula with the more accurate calculation[8] in an example shown in Table 14-3.

We note that the annual default probabilities increase over time since the credit curve is increasing (upwards-sloping), which is a similar effect to that seen in Figure 14-5 for historical default probabilities (e.g., for an A rating). Indeed, the shape of a credit curve plays an important part in determining the distribution of risk-neutral default probability, as we shall see in an example.

Suppose we take three different credit curves: the upwards-sloping one from Table 14-3, a flat curve at 500 bps and an inverted curve with, respectively, spreads of 800 bps, 700 bps, 600 bps, 550 bps and 500 bps. The cumulative default probability curves are shown in

[8] The more accurate calculation assumes constant continuously compounded interest rates of 5% and piecewise constant hazard rates.

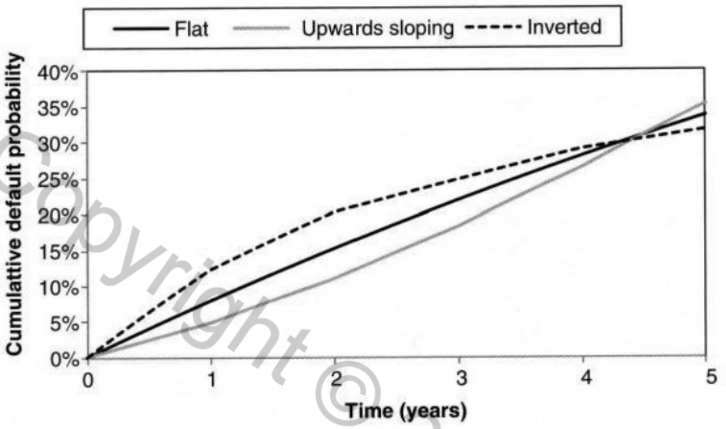


FIGURE 14-8 Cumulative default probabilities for flat, upwards-sloping and inverted curves as described in the text.

In all cases, the 5-year spread is 500 bps and the recovery value is assumed to be 40%.

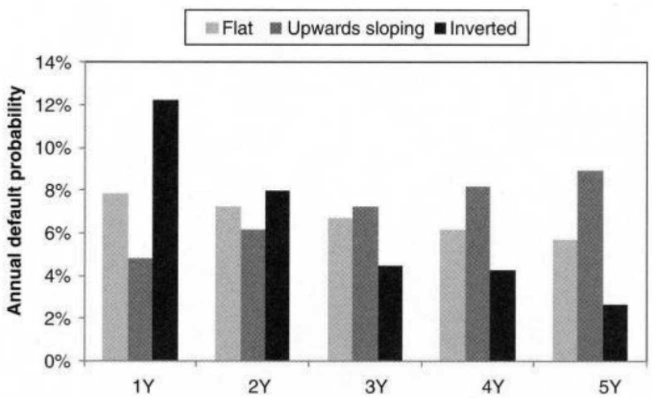


FIGURE 14-9 Annual default probabilities for flat, upwards-sloping and inverted curves as described in the text.

In all cases, the 5-year spread is 500 bps and the recovery value is assumed to be 40%.

Figure 14-8. Note that all have a 5-year credit spread of 500 bps and assumed recovery rates of 40%. The only thing that differs is the shape of the curve. Whilst all curves agree reasonably well on the 5-year cumulative default probability of approximately 33%, the precise shape of the curve up to that point gives very different results. This is seen in Figure 14-9, which shows annual default probabilities for each case. For an upwards-sloping curve, default is less likely in the early years and more likely in the later years, whilst the reverse is seen for an inverted curve. In order to calculate risk-neutral default probabilities properly, in addition to defining the level of the credit curve, it is also important to know the precise curve shape.

Comparison between Real and Risk-Neutral Default Probabilities

The difference between real and risk-neutral default probabilities has been characterised in a number of empirical studies. For example, Giesecke et. al. (2010) use a dataset of bond yields that spans a period of almost 150 years from 1866 to 2008 and find that average credit spreads (averaging across all available bond data) have been about twice as large as realised losses due to default. Studies that are more specific include Fons (1987), the aforementioned work by Altman (1989) and Hull et. al. (2005a). For example, Fons finds that 1-year risk-neutral default probabilities exceed actual realised default rates by approximately 5% (corresponding therefore to the numbers shown in Figure 14-4). The difference between real and risk-neutral default probabilities from Hull et. al. (2005a) is shown in Table 14-4 as a function of credit rating. We see that the difference is large, especially for better quality credits.

There has been much work on understanding the components of credit spreads and their relation to actual default rates and recoveries. See, for example, Collin-Dufresne et. al. (2001) and Downing et. al. (2005). These studies find that the difference between credit spreads and actual default losses is due to:

- the relative illiquidity of corporate bonds requiring a liquidity risk premium;
- the limited upside on holding a bond portfolio, or negative skew in bond returns;[9]
- the non-diversifiable risk of corporate bonds requiring a systemic risk premium.

We do not require here to understand in detail the relationship between credit spreads and historical default

[9] This refers to the fact that the positive return from holding a bond is relatively modest whilst the loss on default can be relatively large.

TABLE 14-4 Comparison between Real and Risk-Neutral Default Probabilities in Basis Points

	Real Default Intensity	Risk-Neutral Default Intensity	Ratio
Aaa	4	67	16.8
Aa	6	78	13.0
A	13	128	9.8
Baa	47	238	5.1
Ba	240	507	2.1
B	749	902	1.2
Caa	1690	2130	1.3

Source: Hull et. al. (2005)

losses, but it is important to appreciate the impact on quantifying and managing counterparty risk. If one does not seek to hedge the default component of counterparty risk, then it is more relevant to consider the real world default probabilities estimated empirically via (for example) historical data.[10] If, on the other hand, one intends to hedge against counterparty defaults, then it is important to consider market credit spreads and associated risk-neutral default probabilities. Clearly, hedging counterparty risk appears to be much more costly than not hedging (Table 14-4).

Most of the empirical evidence compares risk-neutral default probabilities from bond spreads with historical default probabilities. Longstaff et. al. (2005) argue that the CDS market, being more liquid, will not give rise to the liquidity premium represented in Figure 14-4. This would imply that the corresponding ratio between CDS-implied risk-neutral and historical default probabilities would be lower. However, this assumption is not supported by a consistently negative CDS-bond basis. Hence, it seems reasonable that the relationship discussed above is broadly the same when using risk-neutral default probabilities from the CDS market.

[10] Although we note that Basel III regulation is not consistent with the use of real-world default probabilities, irrespective of whether the intention is to hedge or not.

Recovery Rates

In order to estimate risk-neutral default probabilities, we must know the associated recovery rate. Recovery rates refer to the amount that would be recovered in the event of a counterparty defaulting. Common recovery rates are assumed to be a percentage of the notional amount (the exposure). This is in line with the legal right of all creditors to receive a proportion of what they are owed. Recovery rates are sometimes expressed via loss given default (LGD), which is simply one minus the recovery rate (in percentage terms). For example, a low recovery rate of 20% implies a high loss given default of 80%.

Ideally, recovery rates would be derived from market prices. A recovery swap is an agreement between two parties to swap a realised recovery rate (when and if the relevant credit event occurs) with a fixed recovery rate that is specified at the start of the contract. The reference price reflects the fixed recovery such that the recovery swap has zero value initially. Since the swap is issued at a price of zero, if the reference entity does not default in the term of the swap, then the swap expires with no cash flows having taken place. If a default does occur, the fixed recovery payer in the swap will compensate the other party, if the actual recovery is less than the fixed recovery, and vice versa. Since recovery swaps do not trade (except occasionally for distressed credits), we must normally look to historical analysis of recovery rates.

Recovery values, like default probabilities, tend to show a significant variation over time, as illustrated in Figure 14-10. We can see further variation according to

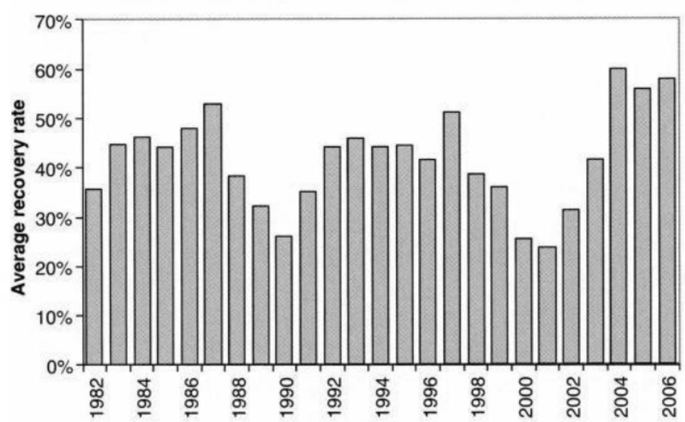


FIGURE 14-10 Average recovery values across all debt seniorities.

Source: Moody's Investors Service (2007)

TABLE 14-5 Recovery Rates by Sector

Industry	Recovery Rate Average
Public utilities	70.5%
Chemicals, petroleum, rubber and plastic products	62.7%
Machinery, instruments and related products	48.7%
Services (business and personal)	46.2%
Food and kindred products	45.3%
Wholesale and retail trade	44.0%
Diversified manufacturing	42.3%
Casino, hotel and recreation	40.2%
Building material, metals and fabricated products	38.8%
Transportation and transportation equipment	38.4%
Communication, broadcasting, movie production, printing and publishing	37.1%
Financial institutions	35.7%
Construction and real-estate	35.3%
General merchandise stores	33.2%
Mining and petroleum drilling	33.0%
Textile and apparel products	31.7%
Wood, paper and leather products	29.8%
Lodging, hospitals and nursing facilities	26.5%
TOTAL	**41.0%**

Source: Altman and Kishore (1996)

variables such as sector (Table 14-5). Recoveries also tend to be negatively correlated with default rates (e.g., see Hamilton et. al., 2001). This negative correlation means that a high default rate will give rise to lower recovery values. Hence, the random nature of default probability and recovery over time coupled to the negative correlation creates strong variability in default losses.

Recovery rates also depend on the seniority of the claim (Table 14-6). Normally, OTC derivatives would rank *pari passu* with senior unsecured debt, which in turn is the reference in most CDS contracts. When the recovery claim for counterparty risk is different, this must be quantified.

A final point on recovery is related to the timing. CDSs are settled quickly following a default and bondholders can settle their bonds in the same process (the CDS auction discussed later). However, OTC derivatives cannot be settled in a timely manner. This is partly due to their bespoke nature and partly due to netting (and collateral), which means that many trades are essentially aggregated into a single claim and cannot be traded individually. The net claim (less any collateral) is then often quite difficult to define for the portfolio of trades. This creates two different recovery values:

- *Settled recovery.* This is the recovery that could be achieved following the credit event; for example, by selling a defaulted bond.