Chapter 2
Credit Risk Measurement in the Context of Basel II

2.1 Banking Supervision and Basel II

During the last decades, there has been a lot of effort spent on improving and extending the regulation of financial institutions. There are several reasons for a regulation of these institutions, which are mostly different from the regulation of other economic sectors. Even if there are some discussions about tendencies of the banking sector to constitute a monopoly as a result of economies of scale and economies of scope, the empirical evidence is rather scarce.[7] A widely accepted argument is that the (unregulated) banking system is unstable. If a bank is threatened by default or the depositors expect a high default risk, this can lead to a bank run, meaning that many depositors could abruptly withdraw their deposits.[8] This behavior is a consequence of the "sequential service constraint", meaning that whether a depositor gets his deposits depends on the position in the waiting queue.[9] The problem is that, as most banks invest the short term deposits in long term projects (term transformation), there is a high risk of illiquidity of the bank, regardless of whether the bank is overindebted or not. Due to incomplete information, depositors of different institutions could also withdraw their deposits, and this domino effect could finally lead to a collapse of the complete banking system. This type of risk is called "systemic risk".[10] Because of the enormous relevance of banks for the complete economy, the state will usually act as a "lender of last resort", especially in the case of big financial institutions ("too big to fail"-phenomenon) instead of accepting a bank's default, which is due to the presence of systemic risk.[11] Against

[7]Cf. Berger et al. (1993, 1999).

[8]Cf. Diamond and Dybvig (1983).

[9]Cf. Greenbaum and Thakor (1995). This is an important difference to securities where the holder is exposed to a price decline instead.

[10]Cf. Saunders (1987) and Hellwig (1995).

[11]The relevance of this phenomenon has been remarkably shown in the ongoing financial crisis. In 2007 and 2008, there have been many examples of bailouts of financial institutions, such as Bear

M. Hibbeln, *Risk Management in Credit Portfolios*, Contributions to Economics,
DOI 10.1007/978-3-7908-2607-4_2,

this background, the state is interested in a regulation of the financial system in order to reduce the probability of bank runs and the systemic risk.[12]

The first German banking supervision was established in 1931 after the default of the Danatbank during the Great Depression. This event came along with a massive withdrawal of deposits and bank runs. As a consequence of the default of the Herstatt Bank in 1974, about 52,000 private customers lost their money. Furthermore, many US American banks, which had currency contracts with the Herstatt Bank, did not get back their receivables. This event led to several additional regulations, including the extension of deposit guarantees and the large exposure rules. Moreover, as a result of this default, the central-bank Governors of the Group of Ten (G10) countries founded the Basel Committee on Banking Supervision in the end of 1974, which had the objective to close gaps in international supervisory coverage. In 1988 the Committee introduced the Basel Capital Accord (*Basel I*), which led to a major harmonization of international banking regulation and minimum capital requirements for banks.[13] According to Basel I, it is required that banks hold equity equal to 8% of their risk weighted assets, which are calculated as a percentage between 0% (e.g. for OECD banks) and 100% (e.g. for corporates) of the credit exposure. The basic principle behind this requirement is that the minimum capital requirement, which also implies a maximum leverage, leads to an acceptable maximum probability of default for every single bank. Thus, this restriction of risk should lead to a stabilization of the banking system. The problem is that these capital rules are hardly risk-sensitive – for example, an investment grade and a speculative grade corporate bond require the identical capital. As a consequence, banks have an incentive to deal with risky credits, especially if the regulatory capital constraint is binding. This incentive stems from the risk-shifting problem, which is relevant for every indebted institution, but increases with leverage. This problem is already present for projects with identical expected pay-offs but as risky investments usually offer higher expected profits, the incentive of risk-shifting is even higher. In addition, the Basel Capital Accord offered the possibility of "regulatory capital arbitrage", which is a result of the missing risk-sensitivity, too. A bank with a small capital buffer could bundle its low-risk assets in asset backed securities and sell them to investors. After this transaction, the bank still has

---

Stearns, Fannie Mae, Freddie Mac, and AIG in the United States or IKB and Hypo Real Estate in Germany. But an even stronger argument for the "too big to fail"-phenomenon is the default of Lehman Brothers in September 2008. Probably due to the global diversification of their creditors, the bank's default was apparently assessed as no systemic risk. But the subsequent financial turmoil including the almost complete dry up of the interbank lending market shows that this was a material misjudgment of the U.S. government; cf. the German Council of Economic Experts (2008), p. 122. This default clearly demonstrates the relevance of the "too big to fail"-phenomenon and the negative consequences if a big financial institution still fails, especially in an unstable market environment.

[12] For a more detailed discussion of banking regulation see Gup (2000) or Hartmann-Wendels et al. (2007), p. 355 ff.

[13] Cf. Phillips and Johnson (2000), p. 5 ff., Hartmann-Wendels et al. (2007), p. 391 ff., Henking et al. (2006), p. 2 ff., and BCBS (2009a).

almost the same degree of risk but free capital, which could be used to invest in new, risky projects. Thus, it is obvious to see that the minimum capital requirements of Basel I do not effectively reduce the risk-taking behavior of banks.

Against this background, in 1999 the Basel Committee on Banking Supervision (BCBS) published the First Consultative Package on a New Basel Capital Accord (*Basel II*) with a more risk-sensitive framework. Finally, in 2004/05 the Committee presented the outcome of its work under the title "Basel II: International Convergence of Capital Measurement and Capital Standards – A Revised Framework" (BCBS 2004c, 2005a). In this context, it is interesting to notice that it was intended to maintain the overall level of regulatory capital.[14] Thus, the purpose of the new capital rules is indeed to achieve better risk-sensitivity. Basel II is based on "three mutually reinforcing pillars, which together should contribute to safety and soundness in the financial system".[15] *Pillar 1* contains the Minimum Capital Requirements, which mainly refer to an adequate capital basis for credit risk, but operational risk and market risk are considered, too. *Pillar 2* is about the Supervisory Review Process. In contrast to Pillar 1, which contains quantitative and qualitative elements, Pillar 2 contains qualitative requirements only. These refer to a proper assessment of individual risks – beyond the demands of Pillar 1 – and sound internal processes in risk management. Important risk types that are not captured by Pillar 1 are concentration risks, which are the object of investigation during this study, interest rate risks, and liquidity risks. *Pillar 3* shall improve the market discipline through an enhanced disclosure by banks, e.g. about the calculation of capital adequacy and risk assessment. The New Basel Capital Accord has been implemented in the European Union in 2006 via the Capital Requirement Directive (CRD). Subsequently, the member states of the European Union transposed the directive into national law. In Germany, the corresponding regulations are basically the "Solvabilitätsverordnung" (SolvV), which refers to the first and third Pillar of Basel II, some changes in the "Kreditwesengesetz" (KWG) and the "Großkredit- und Millionenkreditverordnung" (GroMiKV), as well as the "Mindestanforderungen an das Risikomanagement" (MaRisk), implementing the demands of the Pillar 2. These regulations came into effect on 01-01-2007.

As this study deals with credit risk management, only this type of risk will be considered in the following. In contrast to Basel I, the minimum capital requirements of Basel II take the probabilities of default of the individual credits into consideration. The concrete quantitative requirements are based on a framework that measures the 99.9%-Value at Risk of a portfolio, which is the loss that will not be exceeded with a probability of at least 99.9%. The banks are free to choose the Standardized Approach or the Internal Ratings-Based (IRB) Approach, which mainly differ concerning the use of external ratings vs. internal estimates of the obligors' creditworthiness. Furthermore, for non-retail obligors the IRB Approach is subdivided into the Foundation IRB Approach and the Advanced IRB Approach.

---

[14]BCBS (2001b).

[15]BCBS (2001b), p. 2.

While within the Foundation IRB Approach only the probability of default has to be estimated, banks using the Advanced IRB Approach have to estimate additional parameters, such as the Loss Given Default and the Exposure at Default, which are described in the subsequent Sect. 2.2.1.[16] In this context, it should be noticed that the IRB Approach is not only a regulatory set of rules but the underlying framework often serves as a common fundament in banking practice and for ongoing research in credit risk modeling with several improvements and applications.[17] Against this background, it is useful to have a deeper understanding of the concrete credit risk measurement and credit portfolio modeling as a basis of improving the management of credit risk. Thus, in the following there will be a short introduction on individual risk parameters and risk measures in a credit portfolio context, and a detailed explanation of the framework underlying the IRB Approach.

2.2 Measures of Risk in Credit Portfolios

2.2.1 Risk Parameters and Expected Loss

Before the parameters for the quantification of credit risk are explained, we start with some short comments about the general notation. In the following, stochastic variables are marked with a tilde "$\sim$", e.g. $\tilde{x}$ denotes that x is a random variable. Furthermore, "$\mathbb{E}(\tilde{x})$" stands for the expectation value and "$\mathbb{V}(\tilde{x})$" for the variance of the random variable $\tilde{x}$. Similarly, "$\mathbb{P}(\tilde{x} = a)$" denotes the probability that $\tilde{x}$ takes the value a. The random variable $1_{\{\tilde{x} > a\}}$, which is also called an indicator variable, is defined as

$$1_{\{\tilde{x} > a\}} = \begin{cases} 1 & \text{if } \tilde{x} > a, \\ 0 & \text{if } \tilde{x} \leq a. \end{cases} \tag{2.1}$$

Thus, the indicator variable takes the value one if the event specified in brackets occurs, and zero otherwise. Using this notation, the parameters for the quantification of credit risk can be introduced. The potential loss of a credit is usually expressed as a product of three components: The default indicator variable, the loss given default, and the exposure at default.

---

[16]Details concerning the concrete regulatory requirements and a comparison of these approaches can be found in Heithecker (2007), especially in Sect. 3.

[17]E.g. the underlying one-factor Gaussian copula model with its implied correlation is market standard for pricing CDOs, cf. Burtschell et al. (2007), p. 2, similar to the model of Black and Scholes for options with its implied volatility. Examples for extensions of the standard Gaussian copula model are Andersen and Sidenius (2005a, b) or Laurent and Gregory (2005). Furthermore, several smaller banks use the regulatory capital formulas for their internal capital adequacy assessment process; cf. BCBS (2009b), p. 14.

Firstly, the default event of an obligor is indicated by the *default indicator variable* $1_{\{\tilde{D}\}}$ that takes the value one if the (uncertain) default event $\tilde{D}$ occurs and zero otherwise.[18] The *probability of default* (PD) of an obligor is defined by $\mathbb{P}\left(1_{\{\tilde{D}\}} = 1\right) =: PD$. In context of the Basel Framework, the PD is the probability that an obligor defaults within 1 year.[19] The Basel Committee on Banking Supervision defines a default as follows: "A default is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place:

- The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held).
- The obligor is past due more than 90 days on any material credit obligation to the banking group. Overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstandings".[20]

It is important to notice that beside this definition there exist several other definitions of default[21] so that a credit that is defaulted in Bank A could be treated as non-defaulted in Bank B. But as the definition above has to be implemented at least for regulatory purposes, it can be seen as the conjoint definition of default.

Secondly, the *loss given default* (LGD) gives the fraction of a loan's exposure that cannot be recovered by the bank in the event of default. Besides obligor-specific characteristics the LGD can highly depend on contract-specific characteristics such as the value of collateral and the seniority of the credit obligation. The uncertain LGD is denoted by the random variable $\widetilde{LGD}$, whereas the expected LGD is denoted by $\mathbb{E}(\widetilde{LGD}) =: ELGD$. There also exists a direct link between the loss given default and the so-called recovery rate (RR): $\widetilde{RR} = 1 - \widetilde{LGD}$. Both variables usually take values between 0% and 100% but the LGD can also be higher than 100% as workout costs occur when the bank tries to recover (parts of) the outstanding exposure. If the bank fails to recover the loan, the total loss amount can be higher than the defaulted exposure leading to an effective LGD of more than 100% and to a RR of less than 0%, respectively.

[18]In this study, it is not explicitly differentiated between a default of a single loan or of a firm. In this context, it should be noted that for corporates a defaulting loan is usually associated with a default of the firm; consequently, all other loans of the firm are considered as defaulted, too. Contrary, in retail portfolios the loans are often handled separately; thus, a default of one loan does not imply a default of all other loans of this obligor.

[19]See BCBS (2005a), §§ 285, 331.

[20]BCBS (2005a), § 452. For further details on the definition of default, including a specification of "unlikeliness to pay" see BCBS (2005a), §§ 453–457.

[21]A survey of different definitions of default and their impact on the computed recovery rates can be found in Grunert and Volk (2008).

Thirdly, the *exposure at default* (EAD) consists of the current outstandings (OUT), which are already drawn by the obligor. Furthermore, the obligor could draw a part of the commitments (COMM) leading to an increased EAD. This part is called the credit conversion factor (CCF). Thus, the (uncertain) EAD can be defined as[22]

$$\widetilde{EAD} := OUT + \widetilde{CCF} \cdot COMM \tag{2.2}$$

with $0 \leq \widetilde{CCF} \leq 1$. Despite the fact that the exposure at default is a random variable, it is often associated with "the *expected* gross exposure of the facility upon default of the obligor",[23] that means

$$EAD := OUT + \mathbb{E}\left(\widetilde{CCF}\right) \cdot COMM. \tag{2.3}$$

In this study, the exposure at default is mostly assumed to be deterministic, which leads to identity of the random variable $\widetilde{EAD}$ and the expected value EAD.

Using these three components, we can quantify the loss of a single credit or of a credit portfolio (PF) that consists of n different loans. The loss in absolute values of a single credit $i \in \{1, ..., n\}$ is denoted by $\tilde{L}_{\text{abs},i}$:

$$\tilde{L}_{\text{abs},i} = \widetilde{EAD}_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}. \tag{2.4}$$

Thus, a default of loan i leads to an uncertain loss amount of $\widetilde{EAD}_i \cdot \widetilde{LGD}_i$, which is the fraction LGD of the exposure at default. Similarly, we name the absolute loss of the whole portfolio $\tilde{L}_{\text{abs,PF}}$, which can be calculated as the sum of all individual losses:

$$\tilde{L}_{\text{abs,PF}} = \sum_{i=1}^{n} \tilde{L}_{\text{abs},i} = \sum_{i=1}^{n} \widetilde{EAD}_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}. \tag{2.5}$$

The expected loss $EL_{\text{abs},i}$ of loan i is given by

$$EL_{\text{abs},i} = \mathbb{E}\left(\tilde{L}_{\text{abs},i}\right) = \mathbb{E}\left(\widetilde{EAD}_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}\right) = EAD_i \cdot ELGD_i \cdot PD_i, \tag{2.6}$$

assuming the random variables to be stochastically independent. The expected loss (EL) is also called "standard risk-costs" and the risk premium contained in the

[22] See Bluhm et al. (2003), p. 24 ff.

[23] BCBS (2005a), § 474.

contractual interest rate should at least include this amount.[24] The expected loss of the whole portfolio $EL_{\text{abs,PF}}$ can be calculated as

$$EL_{\text{abs,PF}} = \sum_{i=1}^{n} EL_{\text{abs},i} = \sum_{i=1}^{n} EAD_i \cdot ELGD_i \cdot PD_i. \tag{2.7}$$

Moreover, we differentiate between the absolute and the relative portfolio loss since it is often useful to write the loss in relative terms in analytical credit risk modeling. The relative portfolio loss results when the absolute loss is divided by the total exposure, and will simply be denoted by $\tilde{L}$ in the following:

$$\tilde{L} = \frac{\tilde{L}_{\text{abs,PF}}}{\sum_{j=1}^{n} \widetilde{EAD}_j} = \sum_{i=1}^{n} \frac{\widetilde{EAD}_i}{\sum_{j=1}^{n} \widetilde{EAD}_j} \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} = \sum_{i=1}^{n} \tilde{w}_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}, \tag{2.8}$$

where $\tilde{w}_i := \widetilde{EAD}_i \big/ \sum_{j=1}^{n} \widetilde{EAD}_j$ is the exposure weight of credit i in the portfolio. Using this notation and assuming deterministic exposure weights $w_i = EAD_i \big/ \sum_{j=1}^{n} EAD_j$, the expected relative portfolio loss can be written as

$$EL = \sum_{i=1}^{n} w_i \cdot ELGD_i \cdot PD_i. \tag{2.9}$$

2.2.2 *Value at Risk, Tail Conditional Expectation, and Expected Shortfall*

For an individual loan, the expected loss is the most important risk measure as it significantly influences the contractual interest rate. However, on aggregate portfolio level the quantification of additional risk measures is worthwhile. For instance, it is useful for a bank to get knowledge of the possible portfolio loss in some kind of worst case scenario, which is usually defined with respect to a given confidence level α. Based on this, a bank can determine how much capital is needed to survive such scenarios. There exist several approaches to quantify these capital requirements. Firstly, there are different measures for risk quantification, e.g. the Value at Risk, the Tail Conditional Expectation, and the Expected Shortfall, which will be defined and explained below. Secondly, the capital requirements differ depending on their objective. In Basel II the *regulatory capital* requirement is based on the unexpected loss, which is the difference between the Value at Risk with confidence

[24] Cf. Schroeck (2002), p. 171 f.

level $\alpha = 99.9\%$ and the EL,[25] within a 1-year horizon. Furthermore, banks often internally measure their *economic capital* requirement, which can be defined as the capital level that bank shareholders would choose in absence of capital regulation.[26] The economic capital is usually used for the bank's risk management, the pricing system, the internally defined minimum capital requirement, etc.[27] The internal specification of economic capital can differ from the regulatory capital formula, for instance, regarding the used risk measure, the engine for generating the loss distribution, or the time horizon.[28]

For a definition of the risk measures, a mathematical formulation of *quantiles*, or precisely of the upper quantile q^α and the lower quantile q_α, corresponding to a confidence level α is needed. Given the distribution of a random variable $\tilde{X}$, these quantiles are defined as[29]

$$q_\alpha(\tilde{X}) := \inf\{x \in \mathbb{R} | \mathbb{P}[\tilde{X} \leq x] \geq \alpha\}, \tag{2.10}$$

$$q^\alpha(\tilde{X}) := \inf\{x \in \mathbb{R} | \mathbb{P}[\tilde{X} \leq x] > \alpha\}, \tag{2.11}$$

where $\mathbb{R}$ denotes the set of real numbers. If these definitions are applied to continuous distributions, they lead to the same result. Applied to discrete distributions, the upper quantile can exceed the lower quantile.

The *Value at Risk* (VaR) can be described as "the worst expected loss over a given horizon under normal market conditions at a given confidence level".[30] For an exact formulation, the lower Value at Risk $VaR_\alpha(\tilde{L})$ and the upper Value at Risk $VaR^\alpha(\tilde{L})$ at confidence level α have to be distinguished, which are the quantiles of the loss distribution:[31]

$$VaR_\alpha(\tilde{L}) := q_\alpha(\tilde{L}) = \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} \leq l] \geq \alpha\}, \tag{2.12}$$

$$VaR^\alpha(\tilde{L}) := q^\alpha(\tilde{L}) = \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} \leq l] > \alpha\}. \tag{2.13}$$

---

[25]Sometimes the unexpected loss is defined as $UL = \sqrt{\mathbb{V}(\tilde{L})}$ instead; see e.g. Bluhm et al. (2003), p. 28.

[26]See Elizalde and Repullo (2007).

[27]Cf. Jorion (2001), p. 383 ff.

[28]An extensive overview of current practices in economic capital definition and modeling can be found in BCBS (2009b).

[29]Acerbi and Tasche (2002b), p. 1489.

[30]Jorion (2001), p. xxii. The first known use of the Value at Risk is in the late 1980s by the global research at J.P. Morgan but the first widely publicized appearance of the term was 1993 in the report of the Group of Thirty (G-30), which discussed best risk management practices; cf. Jorion (2001), p. 22.

[31]Cf. Acerbi (2004), p. 155. The slightly different notation results from the definition of l as a loss instead of a profit variable.

For continuous distributions, the definitions are identical and with the definition of a distribution function $F_L(l) = \mathbb{P}(\tilde{L} \leq l)$ the VaR can also be written in terms of the inverse distribution function:

$$\begin{aligned} VaR_\alpha(\tilde{L}) &= \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} \leq l] \geq \alpha\} \\ &= l \text{ with } \mathbb{P}[\tilde{L} \leq l] = \alpha \\ &= l \text{ with } F_L(l) = \alpha \\ &= F_L^{-1}(\alpha). \end{aligned} \tag{2.14}$$

For discrete distributions, the term "Value at Risk" will be referred to the lower Value at Risk $VaR_\alpha(\tilde{L})$ in the following, according to Gordy (2003) and Bluhm et al. (2003), if not indicated differently. Using $\mathbb{P}[\tilde{L} \leq l] = 1 - \mathbb{P}[\tilde{L} > l]$, it follows from (2.12) that

$$\begin{aligned} VaR_\alpha(\tilde{L}) &= \inf\{l \in \mathbb{R} | 1 - \mathbb{P}[\tilde{L} > l] \geq \alpha\} \\ &= \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} > l] \leq 1 - \alpha\}. \end{aligned} \tag{2.15}$$

From this definition the description of the VaR as the minimal loss in the worst $100 \cdot (1 - \alpha)\%$ scenarios can best be seen.[32] Obviously, this risk measure refers to a concrete quantile of a distribution but neglects the possible losses that can occur in the worst $100 \cdot (1 - \alpha)\%$ scenarios.

A risk measure that incorporates these low-probable extreme losses, the so-called tail of the distribution, is the *Tail Conditional Expectation* (TCE). Similar to (2.12) and (2.13) the lower Tail Conditional Expectation $TCE_\alpha(\tilde{L})$ and the upper Tail Conditional Expectation $TCE^\alpha(\tilde{L})$ at confidence level α are defined as the conditional expectations above the corresponding α-quantiles:[33]

$$TCE_\alpha(\tilde{L}) := \mathbb{E}(\tilde{L} | \tilde{L} \geq q_\alpha) = \frac{\mathbb{E}\left(\tilde{L} \cdot 1_{\{\tilde{L} \geq q_\alpha\}}\right)}{\mathbb{P}(\tilde{L} \geq q_\alpha)}, \tag{2.16}$$

$$TCE^\alpha(\tilde{L}) := \mathbb{E}(\tilde{L} | \tilde{L} \geq q^\alpha) = \frac{\mathbb{E}\left(\tilde{L} \cdot 1_{\{\tilde{L} \geq q^\alpha\}}\right)}{\mathbb{P}(\tilde{L} \geq q^\alpha)}. \tag{2.17}$$

Consequently, the TCE is always higher than the corresponding VaR at a given confidence level and can differ for discrete distributions according to the definition

[32] Cf. Acerbi (2004), p. 153.

[33] Acerbi and Tasche (2002b), p. 1490. The loss quantiles $q_\alpha(\tilde{L})$ and $q^\alpha(\tilde{L})$ are abbreviated with q_α and q^α, respectively, to achieve a shorter notation.

of the quantile. For continuous distributions, the upper and lower quantiles are identical and therefore both definitions of TCE equal:

$$TCE^{\alpha}_{\text{cont}}(\tilde{L}) = TCE_{\alpha,\text{cont}}(\tilde{L}) = \mathbb{E}(\tilde{L}|\tilde{L} \geq q_\alpha) = \frac{\mathbb{E}\left(\tilde{L} \cdot 1_{\{\tilde{L} \geq q_\alpha\}}\right)}{\mathbb{P}(\tilde{L} \geq q_\alpha)}$$
$$= \frac{1}{1-\alpha}\mathbb{E}\left(\tilde{L} \cdot 1_{\{\tilde{L} \geq q_\alpha\}}\right). \tag{2.18}$$

Acerbi and Tasche (2002b) introduced a similar risk measure, the *Expected Shortfall* (ES):[34]

$$ES_\alpha(\tilde{L}) := \frac{1}{1-\alpha} \cdot \left(\mathbb{E}\left(\tilde{L} \cdot 1_{\{\tilde{L} \geq q_\alpha\}}\right) - q_\alpha \cdot \left(\mathbb{P}(\tilde{L} \geq q_\alpha) - (1-\alpha)\right)\right). \tag{2.19}$$

In contrast to the VaR and the TCE, the ES only depends on the distribution and the confidence level α but not on the definition of the quantile. Looking at the second term, if the probability that $\tilde{L} \geq q_\alpha$ is higher than $(1-\alpha)$, this fraction has to be subtracted from the conditional expectation. If the probability equals $(1-\alpha)$, as for every continuous distribution, the second term vanishes. In this case, the ES is identical to the TCE. An alternative representation of (2.19) is:[35]

$$ES_\alpha(\tilde{L}) = \frac{1}{1-\alpha}\int_\alpha^1 q^u du. \tag{2.20}$$

The intuition behind the ES and the difference between TCE and ES can be demonstrated with the exemplary probability mass function of a discrete random variable shown in Table 2.1 and the corresponding Fig. 2.1.

In this example, the upper as well as the lower VaR at confidence level $\alpha = 0.95$ is 7%. The corresponding TCE is the expectation conditional on a loss of greater or equal to 7%, which is $7.\bar{4}\%$ in the example. As can be seen in the figure, the probability of the considered events is not equal to 5% but 9%. In contrast to the TCE, for the calculation of the ES, the light grey area is subtracted, which is the

Table 2.1 Loss distribution for an exemplary portfolio

Relative Loss l (in %)	2	4	5	7	8
$\mathbb{P}(\tilde{L} = l)$	80%	10%	1%	5%	4%
$\mathbb{P}(\tilde{L} \leq l)$	80%	90%	91%	96%	100%
$\mathbb{P}(\tilde{L} \geq l)$	100%	20%	10%	9%	4%

[34] Acerbi and Tasche (2002b), p. 1491.

[35] Acerbi and Tasche (2002b), p. 1492.

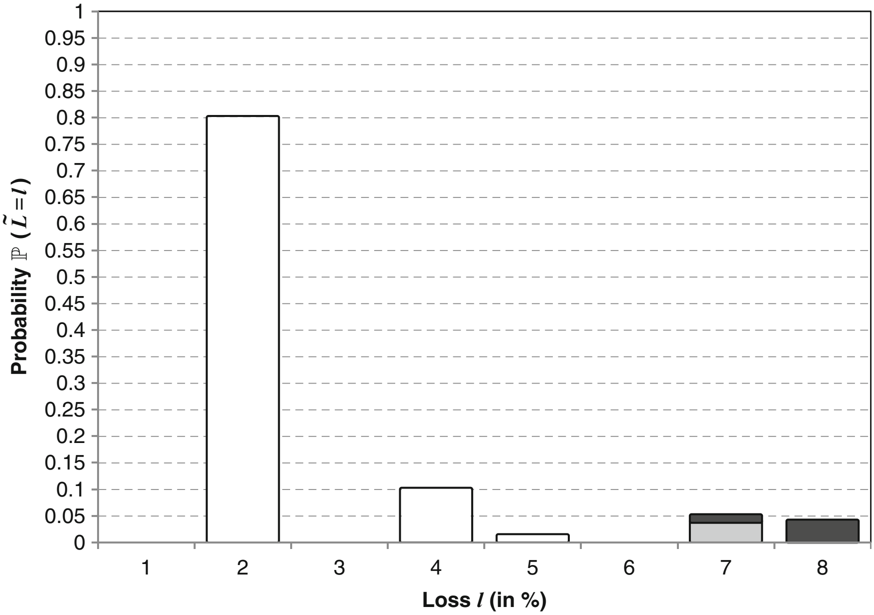

Fig. 2.1 Probability mass function of portfolio losses for an exemplary portfolio

second term of (2.19), and only the dark grey area with a probability of 5% is considered. Thus, the ES is usually higher than the TCE and here we have an ES of 7.8%. Moreover, we can see that the VaR as well as the TCE make a jump if the confidence level is increased from slightly below to slightly above 96%, whereas the ES remains stable because the weight of 7% losses only changes from almost zero to exactly zero.

Subsequently, the calculation of the different risk measures will be demonstrated for the discrete loss distribution of Table 2.1. For this purpose, the confidence levels $\alpha = 0.9$ and $\alpha = 0.95$ are chosen. The upper and lower VaR at these confidence levels are given as

$$VaR_{0.9}(\tilde{L}) = q_{0.9}(\tilde{L}) = \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} \le l] \ge 0.9\} = 4\,\%,$$

$$VaR^{0.9}(\tilde{L}) = q^{0.9}(\tilde{L}) = \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} \le l] > 0.9\} = 5\,\%,$$

$$VaR_{0.95}(\tilde{L}) = q_{0.95}(\tilde{L}) = 7\,\%,$$

$$VaR^{0.95}(\tilde{L}) = q^{0.95}(\tilde{L}) = 7\%.$$

It can be seen that the upper and lower VaR are different if there exists a loss outcome l with $\mathbb{P}(l) > 0$ so that $\mathbb{P}[\tilde{L} \le l] = \alpha$. The same is true for the corresponding TCEs:

$$TCE_{0.9}(\tilde{L}) = \frac{\mathbb{E}\left(\tilde{L}\cdot 1_{\{\tilde{L}\geq q_{0.9}\}}\right)}{\mathbb{P}(\tilde{L}\geq q_{0.9})} = \frac{1}{0.2}(0.1\cdot 4 + 0.01\cdot 5 + 0.05\cdot 7 + 0.04\cdot 8) = 5.6\,\%,$$

$$TCE^{0.9}(\tilde{L}) = \frac{\mathbb{E}\left(\tilde{L}\cdot 1_{\{\tilde{L}\geq q^{0.9}\}}\right)}{\mathbb{P}(\tilde{L}\geq q^{0.9})} = \frac{1}{0.1}(0.01\cdot 5 + 0.05\cdot 7 + 0.04\cdot 8) = 7.2\,\%,$$

$$TCE_{0.95}(\tilde{L}) = \frac{1}{0.09}(0.05\cdot 7 + 0.04\cdot 8) = 7.\bar{4}\,\%,$$
$$TCE^{0.95}(\tilde{L}) = \frac{1}{0.09}(0.05\cdot 7 + 0.04\cdot 8) = 7.\bar{4}\,\%.$$

According to (2.19), there is only one definition of ES, which results in

$$\begin{aligned} ES_{0.9}(\tilde{L}) &= \frac{1}{1-0.9}\left(\mathbb{E}\left[\tilde{L}\cdot 1_{\{\tilde{L}\geq q_{0.9}\}}\right] - q_{0.9}\left[\mathbb{P}\left[\tilde{L}\geq q_{0.9}\right] - (1-0.9)\right]\right)\\ &= \frac{1}{1-0.9}([0.1\cdot 4 + 0.01\cdot 5 + 0.05\cdot 7 + 0.04\cdot 8] - 4\cdot[0.2-0.1]) = 7.2\,\%,\\ ES_{0.95}(\tilde{L}) &= \frac{1}{1-0.95}([0.05\cdot 7 + 0.04\cdot 8] - 7\cdot[0.09-0.05]) = 7.8\,\%. \end{aligned}$$

For demonstration purposes, an ES-definition based on the upper instead of the lower quantile is calculated, too:

$$\begin{aligned} ES^{0.9}(\tilde{L}) &= \frac{1}{1-0.9}\left(\mathbb{E}\left[\tilde{L}\cdot 1_{\{\tilde{L}\geq q^{0.9}\}}\right] - q^{0.9}\left[\mathbb{P}\left[\tilde{L}\geq q^{0.9}\right] - (1-0.9)\right]\right)\\ &= \frac{1}{1-0.9}([0.01\cdot 5 + 0.05\cdot 7 + 0.04\cdot 8] - 5\cdot[0.1-0.1]) = 7.2\,\%,\\ ES^{0.95}(\tilde{L}) &= \frac{1}{1-0.95}([0.05\cdot 7 + 0.04\cdot 8] - 7\cdot[0.09-0.05]) = 7.8\,\%. \end{aligned}$$

It can be seen that the definitions based on the upper as well as on the lower quantile lead to the same result, even if the calculation itself differs for $\alpha = 0.9$.

2.2.3 Coherency of Risk Measures

As demonstrated in Sect. 2.2.2, there exist several measures that could be used for quantifying credit portfolio risk. To identify suitable risk measures, it is reasonable to analyze which mathematical properties should be satisfied by a risk measure to correspond with rational decision making. Based on this, it is possible to evaluate different measures concerning their ability to measure risk in the desired way. Against this background, Artzner et al. (1997, 1999) define a set of four axioms and call the risk measures which satisfy these axioms “coherent”. Some authors even mention that these axioms are the minimum requirements which must be fulfilled by a risk measure and therefore do not distinguish between coherent and non-coherent risk measures but denominate only measures that satisfy these axioms "risk measures".[36]

For a mathematical description of these properties, it is assumed that $\mathcal{G}$ is as set of real-valued random variables (for instance the losses of a set of credits). A function $\rho : \mathcal{G} \to \mathbb{R}$ is called a *coherent risk measure* if the following axioms are satisfied:[37]

(A) *Monotonicity*: $\forall \tilde{L}_1, \tilde{L}_2 \in \mathcal{G}$ with $\tilde{L}_1 \le \tilde{L}_2 \Rightarrow \rho(\tilde{L}_1) \le \rho(\tilde{L}_2)$.

This means that if the losses of portfolio 1 are smaller than the losses of portfolio 2, then the risk of portfolio 1 is smaller than the risk of portfolio 2.

(B) *Subadditivity*: $\forall \tilde{L}_1, \tilde{L}_2 \in \mathcal{G} \Rightarrow \rho(\tilde{L}_1 + \tilde{L}_2) \le \rho(\tilde{L}_1) + \rho(\tilde{L}_2)$.

This axiom reflects the positive effect of diversification. If two portfolios are aggregated, the combined risk should not be higher than the sum of the individual risks. This also means that a merger does not create extra risk. If this axiom is not fulfilled, there is an incentive to reduce the measured risk by asset stripping. Another positive effect is the enabling of a decentralized risk management. If the risk measure ρ is interpreted as the amount of economic capital that is required as a cushion against the portfolio loss, each division of an institution could measure its own risk and could have access to a specified amount of economic capital because the sum of the measured risk or required capital is an upper barrier of the aggregated risk or required capital.

(C) *Positive homogeneity*: $\forall \tilde{L} \in \mathcal{G}, \forall h \in \mathbb{R}^+ \Rightarrow \rho(h \cdot \tilde{L}) = h \cdot \rho(\tilde{L})$.[38]

If a multiple h of an amount is invested into a position, the resulting loss and the required economic capital will be a multiple h of the original loss, too. It is important to notice that this axiom is not necessarily valid for liquidity risk.[39]

---

[36]See e.g. Szegö (2002), p. 1260, and Acerbi and Tasche (2002a), p. 380 f.

[37]Cf. Artzner et al. (1999), p. 209 ff. The definition of the axioms is slightly different from the original set because here the variables $\tilde{L}_1, \tilde{L}_2$ correspond to a portfolio loss instead of a future net worth of a position; see also Bluhm et al. (2003), p. 166. Moreover, it has to be noted that within the axioms of coherency the loss variables $\tilde{L}, \tilde{L}_i$ refer to absolute instead of relative losses.

[38]$\mathbb{R}^+$ denotes all real numbers greater than zero.

[39]The liquidity risk argument is: "If I double an illiquid portfolio, the risk becomes more than double as much!"; see Acerbi and Scandolo (2008), p. 3. Therefore, axiom (B) and (C) are sometimes replaced by a single weaker requirement of *convexity*: $\forall \tilde{L}_1, \tilde{L}_2 \in \mathcal{G}, \forall h \in [0,1] \Rightarrow \rho\left(h \cdot \tilde{L}_1 + (1-h) \cdot \tilde{L}_2\right) \le h \cdot \rho\left(\tilde{L}_1\right) + (1-h) \cdot \rho\left(\tilde{L}_2\right)$; cf. Carr et al. (2001), Frittelli and Rosazza Gianin (2002) or Föllmer and Schied (2002). Acerbi and Scandolo (2008) agree with the statement above but they deny that the coherency axioms are contradicted by this. They argue that the axiom has to be interpreted in terms of portfolio values and not of portfolios. In *liquid* markets the relationship between a portfolio and the value is linear ("if I double the portfolio I double the value"), and therefore there is no difference whether thinking about portfolios or portfolio values. However, in *illiquid* markets the value function is usually non-linear. Based on a proposal of a

(D) *Translation invariance*: $\forall \tilde{L} \in \mathcal{G}, \forall m \in \mathbb{R} \Rightarrow \rho(\tilde{L}+m) = \rho(\tilde{L}) + m$.

If there is an amount m in the portfolio that is lost at the considered horizon with certainty, then the risk is exactly this amount higher than without this position.

In the following, it will be shown that the VaR is not a coherent risk measure as it lacks of subadditivity. The same is true for the TCE if the distribution is discrete.[40] However, the ES satisfies all four axioms and therefore is a (coherent) risk measure.

The *monotonicity* of the VaR directly follows from its definition. If a stochastic variable $\tilde{\varepsilon} \geq 0$ is introduced so that $\tilde{L}_1 + \tilde{\varepsilon} = \tilde{L}_2$, it follows that

$$\begin{aligned} VaR_\alpha(\tilde{L}_1) &= \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L}_1 \leq l] \geq \alpha\} \\ &\leq \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L}_1 \leq l - \tilde{\varepsilon}] \geq \alpha\} \\ &= \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L}_2 \leq l] \geq \alpha\} \\ &= VaR_\alpha(\tilde{L}_2). \end{aligned} \tag{2.21}$$

To show the *positive homogeneity*, a variable $l = h \cdot x$ is introduced so that it follows $\forall \tilde{L} \in \mathcal{G}$ and $\forall h \in \mathbb{R}^+$:

$$\begin{aligned} VaR_\alpha(h \cdot \tilde{L}) &= \inf\{l \in \mathbb{R} | \mathbb{P}[h \cdot \tilde{L} \leq l] \geq \alpha\} \\ &= h \cdot \inf\{x \in \mathbb{R} | \mathbb{P}[h \cdot \tilde{L} \leq h \cdot x] \geq \alpha\} \\ &= h \cdot VaR_\alpha(\tilde{L}). \end{aligned} \tag{2.22}$$

Furthermore, the VaR is *translation invariant* since $\forall \tilde{L} \in \mathcal{G}$ and with $l = x + m$ we obtain:

$$\begin{aligned} VaR_\alpha(\tilde{L} + m) &= \inf\{l \in \mathbb{R} | \mathbb{P}[\tilde{L} + m \leq l] \geq \alpha\} \\ &= \inf\{x \in \mathbb{R} | \mathbb{P}[\tilde{L} + m \leq x + m] \geq \alpha\} + m \\ &= VaR_\alpha(\tilde{L}) + m. \end{aligned} \tag{2.23}$$

The lack of *subadditivity* of the VaR is sufficient to be shown by an example. It is assumed that a loan A and a loan B both have a PD of 6%, an LGD of 100%, and an EAD of 0.5. The VaR at confidence level 90% of each loan is

$$VaR_{0.9}(\tilde{L}_A) = VaR_{0.9}(\tilde{L}_B) = 0. \tag{2.24}$$

formalism for liquidity risk and a proposed non-linear value function, the authors show that liquidity risk is compatible with the axioms of coherency. Further they show that convexity is not a new axiom but a result of the other axioms under their formalism.

[40] Cf. Acerbi and Tasche (2002b), p. 1499, for an example. As the rest of the study focuses on the VaR and the ES, only these risk measures will be analyzed regarding coherency.

If both loans are aggregated into a portfolio, the risk should be smaller or equal to the sum of the individual risks. Assuming that the default events are independent of each other, the probability distribution is given as

$$\begin{aligned} \mathbb{P}\left(\tilde{L}_A + \tilde{L}_B = 0\right) &= (1-0.06)^2 = 88.36\%, \\ \mathbb{P}\left(\tilde{L}_A + \tilde{L}_B = 0.5\right) &= 0.06 \cdot (1-0.06) + (1-0.06) \cdot 0.06 = 11.28\%, \\ \mathbb{P}\left(\tilde{L}_A + \tilde{L}_B = 1\right) &= 0.06^2 = 0.36\%. \end{aligned} \tag{2.25}$$

Thus, the VaR at confidence level 90% of the portfolio is

$$VaR_{0.9}\left(\tilde{L}_A + \tilde{L}_B\right) = 0.5 \tag{2.26}$$

leading to

$$VaR_{0.9}\left(\tilde{L}_A + \tilde{L}_B\right) > VaR_{0.9}\left(\tilde{L}_A\right) + VaR_{0.9}\left(\tilde{L}_B\right). \tag{2.27}$$

This shows that the VaR can be superadditive and thus it is not a coherent risk measure. An important exception is the class of elliptical distributions, e.g. the multivariate normal distribution and the multivariate student's t-distribution, for which the VaR is indeed coherent.[41] As credit risk usually cannot be sufficiently described by elliptical distributions, the lack of coherency can be very critical.

To demonstrate the coherency of ES, it is helpful to use a further representation of (2.19). The purpose is to integrate the second term of (2.19) into the expectation of the first term. Defining a variable $1^{\alpha}_{\{\tilde{L} \geq q_\alpha\}}$ that is

$$1^{\alpha}_{\{\tilde{L} \geq q_\alpha\}} := \begin{cases} 1_{\{\tilde{L} \geq q_\alpha\}} & \text{if } \mathbb{P}\left[\tilde{L} = q_\alpha\right] = 0, \\ 1_{\{\tilde{L} \geq q_\alpha\}} - \frac{\mathbb{P}\left[\tilde{L} \geq q_\alpha\right] - (1-\alpha)}{\mathbb{P}\left[\tilde{L} = q_\alpha\right]} \cdot 1_{\{\tilde{L} = q_\alpha\}} & \text{if } \mathbb{P}\left[\tilde{L} = q_\alpha\right] > 0, \end{cases} \tag{2.28}$$

the ES can be written as[42]

$$ES_\alpha\left(\tilde{L}\right) = \frac{1}{1-\alpha} \cdot \mathbb{E}\left[\tilde{L} \cdot 1^{\alpha}_{\{\tilde{L} \geq q_\alpha\}}\right]. \tag{2.29}$$

For the proof of coherency the following properties will be used:[43]

$$\mathbb{E}\left[1^{\alpha}_{\{\tilde{L} \geq q_\alpha\}}\right] = 1 - \alpha, \tag{2.30}$$

---

[41] Cf. Embrechts et al. (2002). An interesting result is that under the standard assumption of normally distributed returns, Markowitz μ–σ-efficient portfolios are also μ–VaR-efficient.

[42] Cf. Acerbi et al. (2001), p. 8, and Acerbi and Tasche (2002b), p. 1493. For a formal proof see Appendix 2.8.1.

[43] These properties are derived in Appendix 2.8.1, too. See also Acerbi et al. (2001) for a proof based on the ES-definition using upper instead of lower quantiles.

$$1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}} \in [0,\ 1]. \tag{2.31}$$

From definition (2.28) and property (2.31) it can be seen that the variable $1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}$ is not the "normal" indicator function but can also take values between zero and one. Subsequently, the coherency of ES will be shown.[44] The *monotonicity* of the ES can easiest be shown with the integral representation (2.20). It has already been shown that $q_\alpha(\tilde{L}_1) \leq q_\alpha(\tilde{L}_2)$ for $\tilde{L}_1 \leq \tilde{L}_2$ and it can be seen from (2.21) that the same is true for $q^\alpha(\tilde{L}_1) \leq q^\alpha(\tilde{L}_2)$. Therefore, it follows

$$\begin{aligned} ES_\alpha(\tilde{L}_1) &= \frac{1}{1-\alpha}\int_\alpha^1 q^u(\tilde{L}_1)\,du \\ &\leq \frac{1}{1-\alpha}\int_\alpha^1 q^u(\tilde{L}_2)\,du = ES_\alpha(\tilde{L}_2). \end{aligned} \tag{2.32}$$

Using the positive homogeneity of the quantile from (2.22), the ES can shown to be *positive homogeneous* as well:

$$\begin{aligned} ES_\alpha(h\cdot\tilde{L}) &= \frac{1}{1-\alpha}\cdot\mathbb{E}\left[h\cdot\tilde{L}\cdot 1^{\alpha}_{\{h\cdot\tilde{L}\geq q_\alpha(h\cdot\tilde{L})\}}\right] \\ &= \frac{1}{1-\alpha}\cdot\mathbb{E}\left[h\cdot\tilde{L}\cdot 1^{\alpha}_{\{\tilde{L}\geq q_\alpha(\tilde{L})\}}\right] \\ &= h\cdot ES_\alpha(\tilde{L}). \end{aligned} \tag{2.33}$$

The *translation invariance* can be obtained using $\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\right] = 1-\alpha$ (see (2.30)):

$$\begin{aligned} ES_\alpha(\tilde{L}+m) &= \frac{1}{1-\alpha}\cdot\mathbb{E}\left[(\tilde{L}+m)\cdot 1^{\alpha}_{\{(\tilde{L}+m)\geq q_\alpha(\tilde{L}+m)\}}\right] \\ &= \frac{1}{1-\alpha}\cdot\mathbb{E}\left[(\tilde{L}+m)\cdot 1^{\alpha}_{\{\tilde{L}\geq q_\alpha(\tilde{L})\}}\right] \\ &= \frac{1}{1-\alpha}\cdot\mathbb{E}\left[\tilde{L}\cdot 1^{\alpha}_{\{\tilde{L}\geq q_\alpha(\tilde{L})\}}\right] + \frac{m}{1-\alpha}\cdot\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}\geq q_\alpha(\tilde{L})\}}\right] \\ &= ES_\alpha(\tilde{L}) + m. \end{aligned} \tag{2.34}$$

[44] See also Acerbi and Tasche (2002b).

It remains to show the *subadditivity* of the ES. Introducing the random variables $\tilde{L}_1, \tilde{L}_2$ and $\tilde{L}_3 = \tilde{L}_1 + \tilde{L}_2$, the following statement has to be true:

$$ES_\alpha\big(\tilde{L}_1\big) + ES_\alpha\big(\tilde{L}_2\big) - ES_\alpha\big(\tilde{L}_3\big) \geq 0. \tag{2.35}$$

Using representation (2.29) and multiplying by $(1-\alpha)$ leads to

$$\begin{aligned}
&\mathbb{E}\Big[\tilde{L}_1 \cdot 1^\alpha_{\{\tilde{L}_1 \geq q_\alpha(\tilde{L}_1)\}} + \tilde{L}_2 \cdot 1^\alpha_{\{\tilde{L}_2 \geq q_\alpha(\tilde{L}_2)\}} - \tilde{L}_3 \cdot 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big] \\
&\quad = \mathbb{E}\Big[\tilde{L}_1 \cdot \Big(1^\alpha_{\{\tilde{L}_1 \geq q_\alpha(\tilde{L}_1)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) + \tilde{L}_2 \cdot \Big(1^\alpha_{\{\tilde{L}_2 \geq q_\alpha(\tilde{L}_2)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big)\Big].
\end{aligned} \tag{2.36}$$

If the terms in brackets are analyzed, we find that

$$1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}} = \begin{cases} 1 - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}} \geq 0 & \text{if} \quad \tilde{L}_i > q_\alpha\big(\tilde{L}_i\big), \\ 0 - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}} \leq 0 & \text{if} \quad \tilde{L}_i < q_\alpha\big(\tilde{L}_i\big), \end{cases} \tag{2.37}$$

with $i \in [1,\ 2]$, due to the fact that $1^\alpha_{\{\tilde{L} \geq q_\alpha\}} \in [0,\ 1]$. Consequently, we have

$$\begin{aligned}
&\tilde{L}_i \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \\
&\quad \geq q_\alpha\big(\tilde{L}_i\big) \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \quad \text{if} \quad \tilde{L}_i > q_\alpha\big(\tilde{L}_i\big), \\
&\tilde{L}_i \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \\
&\quad \geq q_\alpha\big(\tilde{L}_i\big) \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \quad \text{if} \quad \tilde{L}_i < q_\alpha\big(\tilde{L}_i\big), \\
&\tilde{L}_i \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \\
&\quad = q_\alpha\big(\tilde{L}_i\big) \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \quad \text{if} \quad \tilde{L}_i = q_\alpha\big(\tilde{L}_i\big),
\end{aligned} \tag{2.38}$$

and therefore

$$\tilde{L}_i \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big) \geq q_\alpha\big(\tilde{L}_i\big) \cdot \Big(1^\alpha_{\{\tilde{L}_i \geq q_\alpha(\tilde{L}_i)\}} - 1^\alpha_{\{\tilde{L}_3 \geq q_\alpha(\tilde{L}_3)\}}\Big). \tag{2.39}$$

Using this inequality and again $\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\right]=1-\alpha$ according to (2.30), we find that

$$\begin{aligned}
&\mathbb{E}\left[\tilde{L}_1\cdot\left(1^{\alpha}_{\{\tilde{L}_1\geq q_\alpha(\tilde{L}_1)\}}-1^{\alpha}_{\{\tilde{L}_3\geq q_\alpha(\tilde{L}_3)\}}\right)+\tilde{L}_2\cdot\left(1^{\alpha}_{\{\tilde{L}_2\geq q_\alpha(\tilde{L}_2)\}}-1^{\alpha}_{\{\tilde{L}_3\geq q_\alpha(\tilde{L}_3)\}}\right)\right]\\
&\quad\geq q_\alpha(\tilde{L}_1)\cdot\left(\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}_1\geq q_\alpha(\tilde{L}_1)\}}\right]-\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}_3\geq q_\alpha(\tilde{L}_3)\}}\right]\right)\\
&\quad\quad+q_\alpha(\tilde{L}_2)\cdot\left(\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}_2\geq q_\alpha(\tilde{L}_2)\}}\right]-\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}_3\geq q_\alpha(\tilde{L}_3)\}}\right]\right)\\
&\quad=q_\alpha(\tilde{L}_1)\cdot((1-\alpha)-(1-\alpha))+q_\alpha(\tilde{L}_2)\cdot((1-\alpha)-(1-\alpha))\\
&\quad=0.
\end{aligned}\tag{2.40}$$

Thus, in contrast to the VaR, the ES is subadditive. Since all four axioms are fulfilled, the ES is indeed a coherent risk measure. In addition to the ES, there exist several other coherent risk measures. A class of coherent risk measures is given by the so-called spectral measures of risk with the ES as a special case. This class allows defining a risk-aversion function which leads to different coherent risk measures provided that the risk-aversion function satisfies some conditions presented by Acerbi (2002).[45] However, for the rest of this study the focus will be on the (non-coherent) VaR and the (coherent) ES.

2.2.4 Estimation and Statistical Errors of VaR and ES

Only in minor cases the VaR and the ES will directly be calculated by (2.15) and (2.19), respectively. In real-world applications, the risk measures will mostly be computed via historical simulation or Monte Carlo simulation. In a *historical simulation*, the probability distribution of the loss variable or of several risk factors is assumed to be identical to the empirical distribution of a defined period. Moreover, it is assumed that the realizations are independent of each other. For example, future scenarios will be generated by drawing from $J = 52$ historically observed weekly returns with identical probability. In a *Monte Carlo simulation*, there exists an analytic description of the risk drivers and the dependency between risk drivers and portfolio loss but there is no well-known closed form solution of the probability distribution of the portfolio loss. Thus, a large number J of scenarios can be generated by drawing J independent outcomes of the risk drivers. Using the known dependence structure, J outcomes of the portfolio loss can be computed, which build the simulation-based probability distribution of the portfolio loss.

[45] See also Acerbi (2004), p. 168 ff.

This simulation-based distribution converges towards the exact portfolio distribution as $J \to \infty$.

For a historical simulation as well as for a Monte Carlo simulation, the result is given as a sequence $\{L_j\}_{j=1,\dots,J}$, where each L_j is a realization of the portfolio loss variable $\tilde{L}$. Based on this, the *empirical distribution* is defined as[46]

$$F^{(J)}(l) = \mathbb{P}\big[\tilde{L} \le l\big] = \frac{1}{J} \cdot \sum_{j=1}^{J} 1_{\{L_j \le l\}}. \tag{2.41}$$

For computation of the corresponding VaR and ES, it is useful to introduce the so-called *order statistics* $\{L_{j:J}\}_{j=1,\dots,J}$. Therefore, the sample is sorted into an increasing order such that

$$L_{1:J} \le L_{2:J} \le \dots \le L_{J:J}. \tag{2.42}$$

Now, let $\lfloor x \rfloor$ and $\lceil x \rceil$ denote the floor function and the ceiling function of a real number $x \in \mathbb{R}$:

$$\lfloor x \rfloor = \max\{n \in \mathbb{Z} | n \le x\}, \tag{2.43}$$

$$\lceil x \rceil = \min\{n \in \mathbb{Z} | n \ge x\}, \tag{2.44}$$

where $\mathbb{Z}$ denotes the set of all integers. Then, using the definition of the lower VaR (2.12) and the upper VaR (2.13), the *empirical estimator of VaR* is given as[47]

$$\begin{array}{ll} VaR_\alpha^{(J)}\big(\tilde{L}\big) = VaR^{\alpha(J)}\big(\tilde{L}\big) = L_{\lceil J \cdot \alpha \rceil : J} & \text{if} \quad J \cdot \alpha \notin \mathbb{Z}, \\ \left. \begin{array}{l} VaR_\alpha^{(J)}\big(\tilde{L}\big) = L_{J \cdot \alpha : J} \\ VaR^{\alpha(J)}\big(\tilde{L}\big) = L_{J \cdot \alpha + 1 : J} \end{array} \right\} & \text{if} \quad J \cdot \alpha \in \mathbb{Z}. \end{array} \tag{2.45}$$

This means that except for special cases the VaR is simply given by the $J \cdot \alpha$-th element (rounded up) of the ordered loss sequence. An important characteristic of the empirical estimator is its consistency for large J if the lower VaR equals the upper VaR:

$$\lim_{J \to \infty} VaR_\alpha^{(J)}\big(\tilde{L}\big) = VaR_\alpha\big(\tilde{L}\big) = VaR^\alpha\big(\tilde{L}\big). \tag{2.46}$$

Otherwise the empirical estimators of VaR "flip between the possible values $VaR_\alpha(\tilde{L})$ and $VaR^\alpha(\tilde{L})$".[48]

---

[46]Cf. Acerbi (2004), p. 166.

[47]Cf. also Acerbi (2004), p. 167.

[48]Acerbi (2004), p. 168. This can be illustrated by the "head-or-tail"-example of Acerbi (2004). Let both equiprobable events be related to the loss of $\{-1, 0\}$. The VaRs are given as $VaR_{0.5} = -1$ and $VaR^{0.5} = 0$ but even for large J the 50%-quantile neither converges to –1 nor to 0 but flips between these values.

The *empirical estimator of ES* can be determined with[49]

$$ES_\alpha^{(J)}\left(\tilde{L}\right) = \frac{1}{J \cdot (1-\alpha)} \cdot \left(\sum_{j=\lceil J\cdot\alpha \rceil}^{J} L_{j:J} - (J \cdot \alpha - \lfloor J \cdot \alpha \rfloor) \cdot L_{\lceil J\cdot\alpha \rceil :J} \right). \tag{2.47}$$

In the example of $J = 52$ weekly returns, the 90%-ES can be computed as

$$ES_\alpha^{(52)}\left(\tilde{L}\right) = \frac{1}{5.2} \cdot \left(\sum_{j=47}^{52} L_{j:52} - (46.8 - 46) \cdot L_{47:52} \right). \tag{2.48}$$

This shows that the ES can be interpreted as the average loss in the worst 5.2 scenarios. As can be seen from (2.47), the last term is negligible if J is large. Thus, for historical simulation with a relatively small number of scenarios it is important to consider this term whereas it could be neglected in Monte Carlo simulations since there is typically a very large number of generated scenarios. When $J \cdot \alpha \in \mathbb{Z}$, the empirical estimator simplifies to

$$ES_\alpha^{(J)}\left(\tilde{L}\right) = \frac{1}{J \cdot (1-\alpha)} \cdot \sum_{j=J\cdot\alpha+1}^{J} L_{j:J}. \tag{2.49}$$

Acerbi and Tasche (2002b) showed that the estimator for the ES is consistent for large J:

$$\lim_{J\to\infty} ES_\alpha^{(J)}\left(\tilde{L}\right) = ES_\alpha\left(\tilde{L}\right). \tag{2.50}$$

As shown in the previous sections, the ES has some significant theoretical advantages in comparison with the VaR. But from a practical perspective, the ES is often criticized to be much less robust than the VaR. Consequently, the theoretical advantages of ES could be useless if the number of observations was limited, and thus the VaR would be a much more reliable risk measure than the ES. The standard argument is reproduced by Acerbi (2004) as follows: “VaR does not even try to estimate the leftmost tail events, it simply neglects them altogether, and therefore it is not affected by the statistical uncertainty of rare events. ES on the contrary, being a function of rare events also, has a much larger statistical error”. Against this background, Acerbi (2004) analyzes the statistical errors of VaR

[49] Cf. Acerbi (2004), p. 166 f.

and ES. For continuous distributions of a random variable $\tilde{X}$, the *variances of the estimators* for large J are given as[50]

$$\mathbb{V}\left(VaR_{\alpha}^{(J)}\left(\tilde{X}\right)\right) \overset{J \gg 1}{=} \frac{1}{J} \cdot \frac{\alpha \cdot (1-\alpha)}{f\left(F^{-1}(\alpha)\right)^{2}}, \tag{2.51}$$

$$\mathbb{V}\left(ES_{\alpha}^{(J)}\left(\tilde{X}\right)\right) \overset{J \gg 1}{=} \frac{1}{J \cdot (1-\alpha)^{2}} \cdot \int\limits_{y=0}^{F^{-1}(\alpha)} \int\limits_{z=0}^{F^{-1}(\alpha)} \min(F(y), F(z) - F(y) \cdot F(z)) dz\, dy, \tag{2.52}$$

where F denotes the cumulative distribution function (CDF) of $\tilde{X}$, F^{-1} is the inverse CDF, and $f = dF/dx$ stands for the probability density function (PDF). From (2.51) and (2.52), it can be seen that the estimator of VaR as well as the estimator of ES have the same dependence on the number of trials J. For both estimators, the precision in terms of standard deviation of the demanded statistics can be improved by factor m if the number of trials is increased by factor m^2. However, even if the standard deviations of the estimators are in both cases of order $O(1/\sqrt{J})$,[51] the constant factors could be very different. Therefore, Acerbi (2004) compares the relative error of VaR and of ES for several heavy-tailed probability distributions and confidence levels.[52] He finds that in most cases the relative errors of VaR and ES are very similar. Only in some cases the relative error of ES is at most twice as much as the error of VaR at very high confidence levels. Even if the results of this analysis need not to be true in general, VaR and ES seem to have similar statistical errors and therefore there is no practical burden in implementing the ES instead of the VaR.

2.3 The Unconditional Probability of Default Within the Asset Value Model of Merton

In order to measure the risk of a credit portfolio according to (2.8), it is necessary to specify the stochastic dependence of loan defaults. A widely-used model is the Vasicek model,[53] which is based on the *asset value model* of Merton (1974). In this type of model it is assumed that a firm does not default as a consequence of insufficient liquidity at the moment of repaying a credit because the firm could sell a

[50] Cf. Acerbi (2004), p. 200 f.

[51] The Landau symbol $O(\cdot)$ is defined as in Billingsley (1995), p. 540, A18.

[52] The analyses are performed for lognormal distributions with different volatility parameters and for power law distributions with different shape parameters.

[53] See e.g. Vasicek (1987, 1991, 2002) and Finger (1999, 2001).

part of its assets or it could issue stocks or bonds in order to repay the credit. This can be done as long as the value of liabilities is higher than the value of assets because thenceforward the market participants will not be willing to pay for a security of the firm. Thus, it is assumed that a firm defaults if the asset value $\tilde{A}_T$ is lower than the value of liabilities B payable at time T: $\tilde{A}_T < B$.[54] Consequently, the probability of default is given by

$$PD = \mathbb{P}\left(\tilde{A}_T < B\right). \tag{2.53}$$

The asset value A is modeled as a geometric Brownian motion:[55]

$$dA_t = \mu A_t\, dt + \sigma A_t\, dW_t \quad \text{with} \quad dW_t = \tilde{\varepsilon}\sqrt{dt},\ \tilde{\varepsilon} \sim \mathcal{N}(0,1), \tag{2.54}$$

using the drift rate μ, the volatility σ and the standard Wiener process dW_t.[56] In order to get a closed form solution of the distribution of the asset value at time T, Itô's Lemma is applied to (2.54) leading to[57]

$$dY_t = d\ln A_t = \left(\mu - \frac{1}{2}\sigma^2\right)dt + \sigma\, dW_t. \tag{2.55}$$

This shows that the logarithm of the asset value follows a generalized Wiener process with drift rate $\mu - 1/2\sigma^2$ and variance rate σ^2. As the logarithm of the asset value is normally distributed, the asset value is lognormally distributed. The distribution of the asset value at time T results by integration of (2.55) from $t = 0$ to $t = T$:

$$\begin{aligned}
\ln\left(\frac{\tilde{A}_T}{A_0}\right) &= \ln\tilde{A}_T - \ln A_0 = \int_{t=0}^{T} d\ln A_t \\
&= \int_{t=0}^{T}\left(\mu - \frac{1}{2}\sigma^2\right)dt + \int_{t=0}^{T}\sigma dW_t \\
&= \left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma\left(\tilde{W}_T - W_0\right) \\
&\Leftrightarrow \tilde{A}_T = A_0 \cdot \exp\left[\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma\tilde{W}_T\right],
\end{aligned} \tag{2.56}$$

[54]As can be seen by this expression, the liabilities are assumed to have the structure of a zero coupon bond that has to be paid completely at time T.

[55]A normal distribution with expectation μ and variance σ^2 is indicated by $\mathcal{N}(\mu,\sigma^2)$. Thus, the expression $\tilde{\varepsilon} \sim \mathcal{N}(0,1)$ denotes that $\tilde{\varepsilon}$ follows a standard normal distribution.

[56]For details to the Wiener process see Hull (2006), p. 328 ff.

[57]See Appendix 2.8.2.

using the characteristic of a Wiener process $W_0 = 0$. Using this distribution of the assets at time T from (2.56) and the definition of the Wiener process, the probability of default (2.53) can be calculated:[58]

$$\begin{aligned} PD &= \mathbb{P}\left(\tilde{A}_T < B\right) \\ &= \mathbb{P}\left(\ln\left(\frac{\tilde{A}_T}{A_0}\right) < \ln\left(\frac{B}{A_0}\right)\right) \\ &= \mathbb{P}\left(\left(\mu - \frac{1}{2}\sigma^2\right)T + \sigma\tilde{W}_T < \ln\left(\frac{B}{A_0}\right)\right) \\ &= \mathbb{P}\left(\tilde{\varepsilon}\cdot\sqrt{T} < \frac{\ln\left(\frac{B}{A_0}\right) - \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma}\right) \\ &= \mathbb{P}\left(\tilde{\varepsilon} < -\frac{\ln\left(\frac{A_0}{B}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma\cdot\sqrt{T}}\right) \\ &= \Phi\left(-\frac{\ln\left(\frac{A_0}{B}\right) + \left(\mu - \frac{1}{2}\sigma^2\right)T}{\sigma\cdot\sqrt{T}}\right) \\ &=: \Phi(-\delta). \end{aligned} \tag{2.57}$$

This expression is also known from the Black–Scholes formula of option pricing.[59] The variable δ is called "*distance to default*", as a high value of δ indicates a high equity buffer before a default event can happen. As can be seen in (2.57), the distance to default is higher if the relation of asset to liability value and the drift rate are high and the volatility is low. The problem of asset value models is that the asset value process is not observable and therefore the model cannot easily be calibrated. For firms listed on the stock exchange, the equity values can be observed instead. Therefore, several approaches have been developed for a transformation of equity into asset values.[60]

There also exist several extensions of the asset value model of Merton (1974). Black and Cox (1976) have introduced a *first passage model*, which means that the firm defaults when the asset value is lower than a default barrier for the *first* time and not only at the time of maturity T. In the first passage model of Longstaff and Schwartz (1995) it is assumed that the short-term risk-free interest rate is stochastic, modeled with a Vasicek process, and the risk-free interest rate is correlated with the asset value. Zhou (2001) models the asset return with a jump-diffusion process and thus introduces an additional source of uncertainty leading to empirically more

[59]See Black and Scholes (1973) and Merton (1973).

[60]See for example Bluhm et al. (2003), p. 141 ff. In the documentation of the KMV model (see Crosbie and Bohn 1999) the classical Merton approach is described for solving this problem but according to Bluhm et al. (2003), KMV uses an undisclosed, more complicated algorithm for this task.

plausible results for short-term loans. In addition to the class of asset value models, the probability of default is often determined with *reduced-form models*. In this class, a default is not determined endogenously but it is an exogenous event, and the default time is modeled as the first jump in a jump process. One of the first reduced-form models has been developed by Jarrow and Turnbull (1995).[61] Although the extensions of Merton's asset value model as well as the intensity models usually show a better empirical performance for modeling the PD, it is not necessarily problematic for the validity of the subsequently presented Vasicek model. Even if this model is based on the Merton model, the PD can be determined exogenously with any estimation method as can be seen in the subsequent section.

2.4 The Conditional Probability of Default Within the One-Factor Model of Vasicek

In contrast to the Merton model, the Vasicek model does not focus on the probability of default of a single obligor but quantifies the probability distribution of losses in a loan portfolio. Since the asset value processes and as a consequence the default events cannot be assumed to be independent of each other, a systematic factor is introduced into the model that influences all asset values in a portfolio.[62] As the stochastic interdependence between the firms is modeled by one systematic factor, the model is also called the *Vasicek one-factor model*. The systematic factor is introduced into the model by decomposing the stochastic component of the asset value process from (2.54) or (2.56) into two components that realize at a future point in time T: a systematic part $\tilde{x}$ that influences all firms within the portfolio and a firm-specific (idiosyncratic) part $\tilde{\varepsilon}_i$. Thus, the stochastic component $\tilde{W}_{i,T}$ of each obligor i in $t = T$ can be represented as

$$\tilde{W}_{i,T} = b_i \cdot \tilde{x}_T + c_i \cdot \tilde{\varepsilon}_{i,T}, \tag{2.58}$$

in which $\tilde{x}_T \sim \mathcal{N}(0, T)$ and $\tilde{\varepsilon}_{i,T} \sim \mathcal{N}(0, T)$ are independently and identically normally distributed with mean zero and standard deviation $\sqrt{T}$ for all $i \in \{1, \ldots, n\}$. The degree of the stochastic dependence to the systematic and the idiosyncratic factors is represented by the factor loadings b_i and c_i. In the context of such factor models, the stochastic component $\tilde{W}_i$, mathematically the realization of a standard Wiener process, is usually called the "standardized log-return" of a firm, since this variable results from the logarithm of the asset returns $\ln(\tilde{A}_T / A_0)$ after standardization, see (2.57). For the sake of clarity, the standardized log-returns of the assets

[61] A review of the literature regarding structural and reduced-form models can be found in Duffie and Singleton (2003) and Grundke (2003), p. 15 ff.

[62] Cf. Vasicek (1987).

will be denoted by $\tilde{a}_i$ instead of $\tilde{W}_i$ in the following. Using this notation and choosing a time period of $T = 1$ (e.g. 1 year), (2.58) can be written as

$$\tilde{a}_i = b_i \cdot \tilde{x} + c_i \cdot \tilde{\varepsilon}_i \tag{2.59}$$

with $\tilde{x} \sim \mathcal{N}(0,1)$ and $\tilde{\varepsilon}_i \sim \mathcal{N}(0,1)$. The factor loadings can be written as $b_i = \sqrt{\rho_i}$ and $c_i = \sqrt{1-\rho_i}$, where ρ_i is some constant, as this assures an expectation value of zero and a standard deviation of one of the standardized log-returns $\tilde{a}_i$:

$$\mathbb{E}(\tilde{a}_i) = \mathbb{E}\left(\sqrt{\rho_i} \cdot \tilde{x} + \sqrt{1-\rho_i} \cdot \tilde{\varepsilon}_i\right) = \sqrt{\rho_i} \cdot \mathbb{E}(\tilde{x}) + \sqrt{1-\rho_i} \cdot \mathbb{E}(\tilde{\varepsilon}_i) = 0, \tag{2.60}$$

$$\begin{aligned}\mathbb{V}(\tilde{a}_i) &= \mathbb{V}\left(\sqrt{\rho_i} \cdot \tilde{x} + \sqrt{1-\rho_i} \cdot \tilde{\varepsilon}_i\right) = \rho_i \cdot \mathbb{V}(\tilde{x}) + (1-\rho_i) \cdot \mathbb{V}(\tilde{\varepsilon}_i) \\ &= \rho_i + (1-\rho_i) = 1.\end{aligned} \tag{2.61}$$

In this model, the correlation structure of each firm i is represented by the firm-specific correlation $\sqrt{\rho_i}$ to the common factor.[63] The correlation between the logarithmic asset returns of two firms i, j, which is also called the *asset correlation*, can be expressed as $\sqrt{\rho_i \cdot \rho_j}$ or simply as ρ for the case of a homogeneous correlation structure:

$$\begin{aligned}\rho &= \operatorname{Corr}\left(\ln\left(\frac{\tilde{A}_{i,T}}{A_{i,0}}\right), \ln\left(\frac{\tilde{A}_{j,T}}{A_{j,0}}\right)\right) = \operatorname{Corr}(\tilde{a}_i, \tilde{a}_j) \\ &= \frac{\operatorname{Cov}(\tilde{a}_i, \tilde{a}_j)}{\sqrt{\mathbb{V}(\tilde{a}_i) \cdot \mathbb{V}(\tilde{a}_j)}} = \operatorname{Cov}(\tilde{a}_i, \tilde{a}_j) \\ &= \operatorname{Cov}\left(\sqrt{\rho_i} \cdot \tilde{x} + \sqrt{1-\rho_i} \cdot \tilde{\varepsilon}_i, \sqrt{\rho_j} \cdot \tilde{x} + \sqrt{1-\rho_j} \cdot \tilde{\varepsilon}_j\right) \\ &= \operatorname{Cov}\left(\sqrt{\rho_i} \cdot \tilde{x}, \sqrt{\rho_j} \cdot \tilde{x}\right) = \sqrt{\rho_i} \cdot \sqrt{\rho_j} \cdot \mathbb{V}(\tilde{x}) \\ &= \sqrt{\rho_i \cdot \rho_j}.\end{aligned} \tag{2.62}$$

As already mentioned, within the Vasicek model the probability of default does not have to be computed by the Merton model above but can be used as an exogenously given parameter PD_i.[64] Corresponding to (2.57), an obligor i defaults at $t = T$ when the latent variable $\tilde{a}_i$ falls below a default threshold d_i, which can be characterized by

$$\tilde{a}_i < d_i \quad \Leftrightarrow \quad \sqrt{\rho_i} \cdot \tilde{x} + \sqrt{1-\rho_i} \cdot \tilde{\varepsilon}_i < d_i. \tag{2.63}$$

[63] The factors used in the model are not observable. Therefore, they are also called latent variables.

[64] The probability of default could either be determined by the institution itself or by a rating agency.

Against this background, the threshold d_i can be determined by the exogenous specification of PD_i:[65,66]

$$PD_i = \mathbb{P}\left(1_{\{\tilde{D}_i\}} = 1\right) = \mathbb{P}(\tilde{a}_i < d_i) = \Phi(d_i) \Leftrightarrow d_i = \Phi^{-1}(PD_i). \tag{2.64}$$

Thus, a default event $\tilde{D}_i$ of the firm i can be described by

$$\tilde{D}_i: \quad \tilde{a}_i = \sqrt{\rho_i} \cdot \tilde{x} + \sqrt{1-\rho_i} \cdot \tilde{\varepsilon}_i < \Phi^{-1}(PD_i). \tag{2.65}$$

If the loss distribution of a credit portfolio shall be computed by a Monte Carlo simulation, (2.65) can directly be implemented. In each simulation run the systematic factor as well as the idiosyncratic factors of each obligor are randomly generated. Herewith, the asset return is calculated according to (2.65). If the realization of $\tilde{a}_i$ is less than the threshold given by $\Phi^{-1}(PD_i)$, obligor i defaults. Assuming deterministic LGDs and exposures, the portfolio loss can be determined with formula (2.8) by summing up the exposure weights w_i multiplied by the loss given default LGD_i of each defaulted credit. After repeating this procedure a several thousand times and sorting the losses of the simulation runs, we obtain the portfolio loss distribution. At this point it can be seen that the model of Vasicek does not imply that the PDs are determined on the basis of Merton's asset value model of the previous section. Instead, every estimation method can be used for this purpose and only the dependence structure is specified by the model of Vasicek.

If the loss distribution or some characteristics of the distribution like the VaR or the ES shall be determined analytically, it is helpful to make use of the conditionally independence property of the asset returns. This means that for a given realization of the systematic factor, the asset returns are stochastically independent. Conditional on a realization of the systematic factor $\tilde{x} = x$, the probability of default of each obligor is

$$\begin{aligned}
\mathbb{P}\left(1_{\{\tilde{D}_i\}} = 1 | \tilde{x} = x\right) &= \mathbb{P}(\tilde{a}_i < d_i | \tilde{x} = x) \\
&= \mathbb{P}\left(\sqrt{\rho_i} \cdot \tilde{x} + \sqrt{1-\rho_i} \cdot \tilde{\varepsilon}_i < \Phi^{-1}(PD_i) | \tilde{x} = x\right) \\
&= \mathbb{P}\left(\tilde{\varepsilon}_i < \frac{\Phi^{-1}(PD_i) - \sqrt{\rho_i} \cdot x}{\sqrt{1-\rho_i}}\right) \\
&= \Phi\left(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho_i} \cdot x}{\sqrt{1-\rho_i}}\right) =: p_i(x).
\end{aligned} \tag{2.66}$$

[65] The function $\Phi^{-1}(\cdot)$ stands for the inverse standard normal CDF.

[66] If the probability of default is determined by the asset value model, the default threshold d_i equals the negative distance to default $-\delta$, see (2.57).

This *conditional probability of default* $p_i(x)$ is the PD that would be assigned if the realization of the systematic factor at the horizon was known. By contrast, the unconditional probability of default reflects all information that is currently available, which means that the systematic factor is a random variable and therefore unknown. The unconditional PD equals the average value of the conditional PD across all possible realizations of the systematic factor.[67] This can be shown using the law of iterated expectations:[68]

$$\begin{aligned}\mathbb{E}(p_i(\tilde{x})) &= \mathbb{E}\left(\mathbb{P}\left(1_{\{\tilde{D}_i\}} = 1|\tilde{x}\right)\right) = \mathbb{E}\left(\mathbb{E}\left(1_{\{\tilde{D}_i\}}|\tilde{x}\right)\right) \\ &= \mathbb{E}\left(1_{\{\tilde{D}_i\}}\right) = \mathbb{P}\left(1_{\{\tilde{D}_i\}} = 1\right) = PD_i.\end{aligned} \tag{2.67}$$

Formula (2.66) for the conditional probability of default is sometimes called the *Vasicek formula* and is also used within the Basel framework. Details will be described in Sect. 2.7.

2.5 Measuring Credit Risk in Homogeneous Portfolios with the Vasicek Model

In order to achieve an analytical solution of the loss distribution, it is helpful to assume that the credit portfolio is homogeneous. In a homogeneous portfolio, all credits have the same PD, an identical (deterministic) LGD, the same EAD, and an identical asset correlation:[69]

$$PD_i = PD,\ LGD_i = LGD,\ EAD_i = EAD, \text{ and } \rho_i = \rho \quad \forall i = 1, \ldots, n. \tag{2.68}$$

In (sub-)portfolios where the credits have similar exposures and similar risk characteristics the assumption of homogeneity should not be critical and lead to a good approximation of the loss distribution. Candidates for application of such a simplification are retail portfolios and in some cases portfolios of smaller banks.[70] In a homogeneous portfolio, a default of k credits leads to a relative loss of

$$l = \frac{k \cdot EAD \cdot LGD}{n \cdot EAD} = \frac{k}{n} \cdot LGD. \tag{2.69}$$

[67] Cf. Gordy (2003), p. 203.

[68] Cf. Franke et al. (2004), p. 41.

[69] This section is based on Vasicek (1987).

[70] Cf. Bluhm et al. (2003), p. 60.

As the defaults are exchangeable, this loss results for any k defaults. The probability of this event is

$$\mathbb{P}\left(\sum_{i=1}^{n} 1_{\{\tilde{D}_i\}} = k\right) = \underbrace{\binom{n}{k}}_{*} \cdot \mathbb{P}\left(\underbrace{\tilde{A}_{1,T} < B_1, \ldots, \tilde{A}_{k,T} < B_k}_{**}, \underbrace{\tilde{A}_{k+1,T} \geq B_{k+1}, \ldots, \tilde{A}_{n,T} \geq B_n}_{***}\right). \tag{2.70}$$

The expression $\tilde{A}_{i,T} < B_i$ indicates a default of firm i.[71] Therefore, the term (**) refers to a default of the first k credits, whereas the other n–k credits (***) do not default. The binomial coefficient (*) represents the number of possible combinations of k defaults out of n credits. Using the conditional independence property of Sect. 2.4, the probability of having k defaults can easily be computed within the one-factor model:[72]

$$\begin{aligned}
P_k &= \mathbb{P}\left(\sum_{i=1}^{n} 1_{\{\tilde{D}_i\}} = k\right) \\
&= \binom{n}{k} \cdot \mathbb{P}\left(\tilde{A}_{1,T} < B_1, \ldots, \tilde{A}_{k,T} < B_k, \tilde{A}_{k+1,T} \geq B_{k+1}, \ldots, \tilde{A}_{n,T} \geq B_n\right) \\
&= \binom{n}{k} \cdot \int_{x=-\infty}^{\infty} \mathbb{P}\left(\tilde{A}_{1,T} < B_1, \ldots, \tilde{A}_{k,T} < B_k, \tilde{A}_{k+1,T} \geq B_{k+1}, \ldots, \tilde{A}_{n,T} \geq B_n \middle| \tilde{x} = x\right) d\Phi(x) \\
&= \binom{n}{k} \cdot \int_{x=-\infty}^{\infty} \mathbb{P}\left(\tilde{\varepsilon}_1 < \frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot x}{\sqrt{1-\rho}}, \ldots, \tilde{\varepsilon}_k < \frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot x}{\sqrt{1-\rho}},\right. \\
&\qquad \left. \tilde{\varepsilon}_{k+1} \geq \frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot x}{\sqrt{1-\rho}}, \ldots, \tilde{\varepsilon}_n \geq \frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot x}{\sqrt{1-\rho}}\right) d\Phi(x) \\
&= \int_{x=-\infty}^{\infty} \binom{n}{k} \cdot (p(x))^k \cdot (1 - p(x))^{n-k} d\Phi(x).
\end{aligned} \tag{2.71}$$

This is also known as the *Vasicek binomial model* since the number of defaults (and the gross loss rate) of the portfolio is binomially distributed with probability $p(x)$ for a realization of the systematic factor $\tilde{x} = x$:[73]

---

[71]Cf. Sect. 2.3

[72]The second step is performed by using the Bayes' theorem for continuous distributions, cf. Appendix 2.8.3, and the standard normal distribution of the systematic factor.

[73]The notation $\mathcal{B}(n,p)$ indicates a binomial distribution with parameters n and p.

$$\left(\sum_{i=1}^{n} 1_{\{\tilde{D}_i\}}|x\right) \sim \mathcal{B}(n, p(x)). \tag{2.72}$$

Hence, the conditional probability of k defaults equals

$$\mathbb{P}\left(\sum_{i=1}^{n} 1_{\{\tilde{D}_i\}} = k|\tilde{x} = x\right) = \binom{n}{k} \cdot (p(x))^k \cdot (1 - p(x))^{n-k}, \tag{2.73}$$

which is the integrand of (2.71).[74]

Due to the homogeneity of exposures, the corresponding loss distribution function is given as[75]

$$\begin{aligned} F^{(n)}(l) = \mathbb{P}\big(\tilde{L} \leq l\big) &= \mathbb{P}\left(\frac{1}{n} \cdot LGD \cdot \sum_{i=1}^{n} 1_{\{\tilde{D}_i\}} \leq l\right) \\ &= \mathbb{P}\left(\sum_{i=1}^{n} 1_{\{\tilde{D}_i\}} \leq \frac{l \cdot n}{LGD}\right) = \sum_{k=0}^{\lfloor l \cdot n/LGD \rfloor} P_k. \end{aligned} \tag{2.74}$$

With (2.71) and (2.74), the distribution can be computed via numerical integration; thus, in the case of homogeneous portfolios, there is no need for a Monte Carlo simulation. Furthermore, applying definition (2.15) and (2.19), the risk measures VaR and ES within the Vasicek binomial model can be computed, which will be named $VaR^{(n)}(l)$ and $ES^{(n)}(l)$, respectively, leading to

$$VaR_\alpha^{(n)}\big(\tilde{L}\big) = \inf\left\{ l \in \mathbb{R} | \mathbb{P}\big[\tilde{L} \leq l\big] = \sum_{k=0}^{\lfloor l \cdot n/LGD \rfloor} P_k \geq \alpha \right\}, \tag{2.75}$$

$$ES_\alpha^{(n)}\big(\tilde{L}\big) = \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L} \cdot 1_{\{\tilde{L} \geq VaR_\alpha^{(n)}\}}\right] - VaR_\alpha^{(n)}\left[\mathbb{P}\left[\tilde{L} \geq VaR_\alpha^{(n)}\right] - (1-\alpha)\right]\right). \tag{2.76}$$

If it is assumed that the portfolio consists of an infinite number of obligors,[76] an easy-to-handle closed form solution of the loss distribution and the probability

[74] See also Gordy and Heitfield (2000).

[75] The symbolism $\lfloor x \rfloor$ is defined as in (2.43).

[76] In this case, the homogeneous portfolio is called “infinitely fine grained”. See also Sect. 2.6 for further details.

density function can be achieved. According to Vasicek (1991), the resulting *limit distribution* is[77]

$$\begin{aligned}F^{(\infty)}(l) &= \lim_{n\to\infty} F^{(n)}(l)\\ &= \Phi\left(\frac{1}{\sqrt{\rho}}\cdot\left(\sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{l}{LGD}\right)-\Phi^{-1}(PD)\right)\right)\end{aligned} \tag{2.77}$$

and the corresponding probability density function equals

$$\begin{aligned}f^{(\infty)}(l) =& \sqrt{\frac{1-\rho}{\rho}}\\ &\cdot\exp\left(-\frac{1}{2\rho}\cdot\left[\sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{l}{LGD}\right)-\Phi^{-1}(PD)\right]^2+\frac{1}{2}\left[\Phi^{-1}\left(\frac{l}{LGD}\right)\right]^2\right).\end{aligned} \tag{2.78}$$

Both functions are visualized in Fig. 2.2 for the parameter setting $PD = 5\%$, $\rho = 20\%$, and $LGD = 100\%$. Obviously, the probability density function is

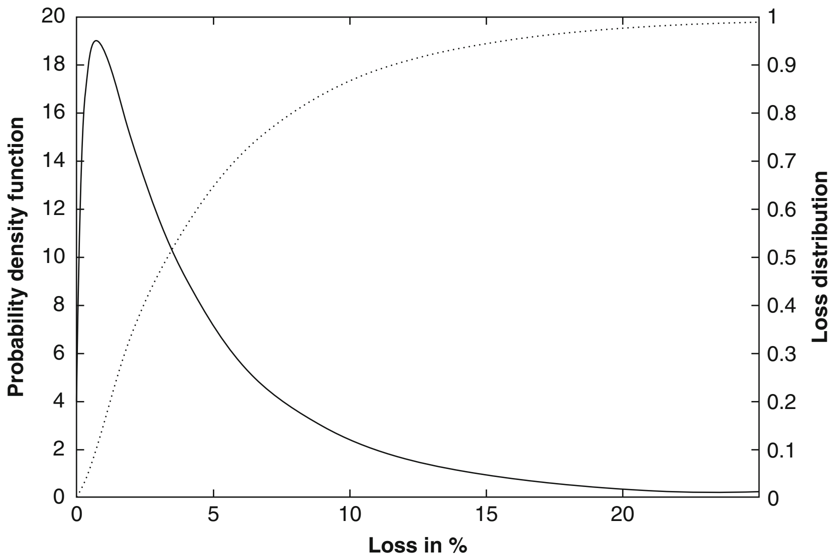

Fig. 2.2 Limiting loss distribution of Vasicek (1991)

[77]See Appendix 2.8.4.

right-skewed and the function has so-called "fat tails". Thus, the kurtosis of loss distributions is typically much higher than the kurtosis of a standard normal distribution. These characteristics reflect the relatively high probability of suffering losses that are several times higher than the expected loss.

With this resulting limit distribution, it is possible to quickly approximate the loss distribution of large subportfolios with similar risk characteristics with high accuracy. This could especially be done for subsegments of a bank's retail portfolio. Furthermore, as the distribution only depends on the PD, the LGD, and the correlation parameter, the complexity of model calibration is relatively low. Based on the loss distribution (2.77) the VaR and the ES can be computed in closed form, too:[78]

$$VaR_{\alpha}^{(\infty)}\left(\tilde{L}\right) = \Phi\left(\frac{\Phi^{-1}(PD) + \sqrt{\rho}\cdot\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right)\cdot LGD, \tag{2.79}$$

$$ES_{\alpha}^{(\infty)}\left(\tilde{L}\right) = \frac{1}{1-\alpha}\cdot\Phi_2\left(\Phi^{-1}(PD), -\Phi^{-1}(\alpha), \sqrt{\rho}\right)\cdot LGD, \tag{2.80}$$

where $\Phi_2(\cdot)$ stands for the bivariate cumulative normal distribution function. This function is defined as

$$\Phi_2\left(x, y, \varrho^2\right) := \mathbb{P}\left(\tilde{X} \le x, \tilde{Y} \le y\right) = \int\limits_{u=-\infty}^{x}\int\limits_{v=-\infty}^{y}\varphi_2(u, v)dv\,du, \tag{2.81}$$

where $\tilde{X}$, $\tilde{Y}$ are standard normal distributed random variables, which have a correlation of ϱ. The joint density function φ_2 of the bivariate standard normal distribution is defined as[79]

$$\varphi_2(u, v) := \frac{1}{2\pi\sqrt{1-\varrho^2}}\cdot\exp\left(-\frac{1}{2}\frac{u^2 - 2\varrho uv + v^2}{1-\varrho^2}\right). \tag{2.82}$$

2.6 Measuring Credit Risk in Heterogeneous Portfolios with the ASRF Model of Gordy

In order to achieve analytical tractability of a model that can be used for risk quantification in heterogeneous portfolios, the so-called *Asymptotic Single Risk Factor (ASRF) framework* has been developed by Gordy (2003).[80] In this framework it is assumed that

[78]See Appendix 2.8.5.

[79]Cf. Bronshtein et al. (2007), p. 779 f., especially (16.156).

[80]See also Bank and Lawrenz (2003).

(A) The portfolio is *infinitely fine-grained* and
(B) Only a *single systematic risk* factor influences the credit risk of all loans in the portfolio

Assumption (A) refers to the granularity of a portfolio that describes the impact of a single credit to the overall portfolio. In a portfolio that consists of a small number of borrowers – a *coarse-grained portfolio* – there is a relatively high impact of the firm-specific, idiosyncratic risk component. A portfolio with a high degree of name concentration is also called a “lumpy” credit portfolio. In contrast, the idiosyncratic risk vanishes in the limiting case of infinite granularity and the risk is solely a result of the uncertainty about the systematic risk factor,[81] as will be shown in the following. A portfolio is “*infinitely granular*” or “*asymptotic*” if it consists of a nearly infinite number of credits $(n \to \infty)$ with each credit having a deterministic exposure weight of negligible size. Concretely, the following conditions have to be fulfilled:[82]

$$\lim_{n\to\infty} \sum_{i=1}^{n} EAD_i = \infty, \tag{2.83}$$

$$\sum_{n=1}^{\infty} \left(\frac{EAD_n}{\sum\limits_{j=1}^{n} EAD_j} \right)^2 < \infty. \tag{2.84}$$

Furthermore, it is assumed that all dependencies across credit events can be expressed by a set of systematic risk factors $\tilde{x}$ so that the credit events are mutually independent conditional on $\tilde{x}$.[83] This not only refers to the assumption of conditionally independent defaults but also to conditional independence of LGDs and especially of the products $(\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}})$. These conditions are necessary for the applicability of the strong law of large numbers. As shown in Appendix 2.8.7, these conditions assure that the portfolio loss (almost surely) equals its conditional expectation:

$$\mathbb{P}\left(\lim_{n\to\infty} \left[\tilde{L} - \mathbb{E}\left(\tilde{L}|\tilde{x}\right) \right] = 0 \right) = 1, \tag{2.85}$$

[81]Cf. BCBS (2001a), p. 89, § 422. This effect could also be found for the limiting distribution of the Vasicek binomial model, see Sect. 2.5.

[82]Cf. Bluhm et al. (2003), p. 87 ff.

[83]Assumption (B), the existence of only a single systematic risk factor, is not needed at this stage.

which is usually much easier to calculate than the unconditional loss distribution.[84] As demonstrated in Appendix 2.8.8, (2.83) and (2.84) also assure that[85]

$$\lim_{n\to\infty} \sum_{i=1}^{n} {w_i}^2 = 0. \tag{2.86}$$

Thus, the weight of each exposure must be negligible. This formulation is directly related to the Herfindahl–Hirschmann Index (HHI), a common measure for indicating the degree of concentration:[86]

$$HHI = \sum_{i=1}^{n} {w_i}^2 = \frac{1}{n^*}. \tag{2.87}$$

In contrast to the actual number of credits n, the variable n^* is the so-called "effective number" of credits. In a homogeneous portfolio, which has the least possible exposure concentration for a given number of credits, n and n^* are identical. Hence, n^* can be interpreted as the number of credits in a homogeneous portfolio with the equivalent degree of name concentration risk. (2.86) can therewith be formulated as

$$\lim_{n\to\infty} \sum_{i=1}^{n} {w_i}^2 = \lim_{n\to\infty} \frac{1}{n^*} = 0, \tag{2.88}$$

which shows that it is not enough that the *actual* number of credits goes to infinity but the *effective* number of credits must go to infinity.

Using property (2.85) the VaR can be written as[87]

$$\lim_{n\to\infty} VaR_\alpha\left(\tilde{L}\right) = VaR_\alpha\left(\mathbb{E}\left[\tilde{L}|\tilde{x}\right]\right). \tag{2.89}$$

Additionally, Gordy (2003) has introduced assumption (B), which states that there is only a single risk factor that influences the credit risk of all loans. Thus, it is assumed that there exist no sector-specific risk factors such as industry-specific or

[84] For ease of notation, the convergence of a sequence X_n towards X with probability one is indicated by $\lim_{n\to\infty} X_n = X$ instead of $\mathbb{P}\left(\lim_{n\to\infty} X_n = X\right) = 1$ in the following.

[85] This is the result of Kronecker's Lemma, see Appendix 2.8.8, which is also needed to proof the strong law of large numbers presented in Appendix 2.8.7. This condition has also been formulated by Vasicek (2002), p. 160.

[86] See BCBS (2001a), p. 97, § 459 and Gordy (2003). The HHI was used in this earlier version of the Basel framework for mapping a heterogeneous portfolio into a comparable homogeneous portfolio.

[87] See Gordy (2003), p. 206 ff.

geographical risk factors and consequently no concentrations in specific sectors. If assumptions (A) and (B) are fulfilled, the following identity holds:[88]

$$VaR_\alpha\left(\mathbb{E}\left[\tilde{L}|\tilde{x}\right]\right) = \mathbb{E}\left(\tilde{L}|\tilde{x} = VaR_{1-\alpha}(\tilde{x})\right). \tag{2.90}$$

This leads to the important proposition

$$VaR_\alpha^{(\mathrm{ASRF})} = \lim_{n\to\infty} VaR_\alpha\left(\tilde{L}\right) = VaR_\alpha\left(\mathbb{E}\left[\tilde{L}|\tilde{x}\right]\right) = \mathbb{E}\left(\tilde{L}|\tilde{x} = VaR_{1-\alpha}(\tilde{x})\right). \tag{2.91}$$

As a result of the conditional independence of all credit events, this proposition can be written as

$$\begin{aligned} VaR_\alpha^{(\mathrm{ASRF})} &= \mathbb{E}\left(\sum_{i=1}^{n} w_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} |\tilde{x} = VaR_{1-\alpha}(\tilde{x})\right) \\ &= \sum_{i=1}^{n} \mathbb{E}\left(w_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} |\tilde{x} = VaR_{1-\alpha}(\tilde{x})\right). \end{aligned} \tag{2.92}$$

It is obvious that the risk contribution of a single credit is equal to its conditional expected loss and is therefore constant, regardless of the concrete portfolio to which the credit is added. This characteristic is also called *portfolio-invariance*. This can be explained by the fact that each individual claim does not cause any (further) diversification effect, since the portfolio has already reached the highest possible degree of diversification. A further important implication is that the VaR of a portfolio is exactly additive because the expected value is exactly additive as well. Consequently, the axiom of subadditivity holds and the VaR is a coherent risk measure under the assumptions described above.[89]

The corresponding expression for the risk measure ES is[90]

$$\lim_{n\to\infty} ES_\alpha\left(\tilde{L}\right) = ES_\alpha\left(\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right) \tag{2.93}$$

leading to

$$ES_a^{(\mathrm{ASRF})}\left(\tilde{L}\right) = ES_\alpha\left(\sum_{i=1}^{n} \mathbb{E}\left(w_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} |\tilde{x}\right)\right). \tag{2.94}$$

[88]See Appendix 2.8.9. The slightly different result concerning the confidence level results from a different definition of the systematic factor. Gordy (2003) assumes that the expected loss is monotonously *increasing* in x, whereas here it is assumed that the expected loss is monotonously *decreasing* in x. In other words, large values of x indicate a good economic condition in this setting.

[89]Cf. Sect. 2.2.2.

[90]See Appendix 2.8.10.

Although the equivalent to (2.91) cannot be formulated for the ES in general form, many specified single-factor models still allow to determine the ES analytically.[91]

2.7 Measuring Credit Risk Within the IRB Approach of Basel II

The IRB Approach of Basel II is based on both the ASRF framework of Gordy (2003) and the conditional probability of default resulting from Vasicek (1987). Under the assumptions of the ASRF framework, it has been shown that the VaR is given as

$$VaR_{\alpha}^{(\mathrm{ASRF})}\left(\tilde{L}\right) = \sum_{i=1}^{n} \mathbb{E}\left(w_i \cdot \widetilde{LGD}_i \cdot 1_{\left\{\tilde{D}_i\right\}} | \tilde{x} = VaR_{1-\alpha}(\tilde{x})\right). \tag{2.95}$$

The confidence level is chosen as $\alpha = 0.999$ in the Basel framework.[92] Furthermore, the conditional probability of default is specified to

$$\mathbb{P}\left(1_{\left\{\tilde{D}_i\right\}} = 1 | \tilde{x} = x\right) = \mathbb{E}\left(1_{\left\{\tilde{D}_i\right\}} | \tilde{x} = x\right) = \Phi\left(\frac{\Phi^{-1}(PD_i) - \sqrt{\rho_i} \cdot x}{\sqrt{1-\rho_i}}\right), \tag{2.96}$$

which is a result of the Vasicek one-factor model. Recalling the standard normal distribution of the systematic factor, the VaR can be written as

$$\begin{aligned} VaR_{0.999}^{(\mathrm{Basel})}\left(\tilde{L}\right) &= \sum_{i=1}^{n} \mathbb{E}\left(w_i \cdot \widetilde{LGD}_i \cdot 1_{\left\{\tilde{D}_i\right\}} | \tilde{x} = VaR_{0.001}(\tilde{x})\right) \\ &= \sum_{i=1}^{n} w_i \cdot \mathbb{E}\left(\widetilde{LGD}_i | \tilde{x} = \Phi^{-1}(0.001)\right) \cdot \mathbb{E}\left(1_{\left\{\tilde{D}_i\right\}} | \tilde{x} = \Phi^{-1}(0.001)\right) \\ &= \sum_{i=1}^{n} w_i \cdot \mathbb{E}\left(\widetilde{LGD}_i | \tilde{x} = -\Phi^{-1}(0.999)\right) \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i} \cdot \Phi^{-1}(0.999)}{\sqrt{1-\rho_i}}\right). \end{aligned} \tag{2.97}$$

This is the core element of the Basel II framework, even if there are some minor differences between the formula above and the concrete capital requirements. These differences are:

[91]Cf. Gordy (2003), p. 219.

[92]From the second to the third consultative document of the Basel framework, the confidence level was changed from $\alpha = 0.995$ to $\alpha = 0.999$; cf. BCBS (2001a, 2003a).

- The capital requirements are only applied to the *Unexpected Loss* (UL), which is the difference of VaR and EL. This is due to the fact that the expected loss is already accounted for in the provisions. As the loan loss provisioning reduces the equity, a capital requirement which includes the expected loss would require this capital amount twice.[93]
- The LGD-specific term of (2.97) shows that the expected LGD under the specified conditions of a VaR scenario is needed. The regulatory formula simply uses the notation "LGD" in the VaR term as well as in the expected loss term. However, this does not mean that the expected LGD has to be inserted. If an institution uses own LGD estimates, these have to "reflect economic downturn conditions where necessary to capture the relevant risks".[94] This LGD is also called "*Downturn LGD*" (DLGD). A background note on LGD quantification clarifies that the downturn LGD is at least in principle meant in terms of the conditional LGD of (2.97). But as a concrete quantification and validation of downturn LGDs in the sense above is found to be "not operationally feasible given the current state of practice in this area", there is no regulatory function that transforms the unconditional into a conditional LGD and also no explicit demand for LGD quantification in a 99.9% scenario.[95]
- The PD in the formula above refers to the 1-year probability of default. In practice, many loans have an effective maturity M_i that can substantially differ from 1 year, especially towards longer maturities. As a long-term loan is usually considered as more risky than a short-term loan, this shall also be reflected in the capital requirement. Therefore a so-called *Maturity Adjustment* is implemented as a factor in the Basel II capital rules.[96]
- The overall level of minimum capital requirements of the model above is calibrated to a regulatory desired magnitude by introducing a *Scaling Factor* (SF), which has to be multiplied to the result of the model itself. This factor is set

---

[93]Because of this argument, the former version of the capital rules, which had the VaR and not the UL as capital requirement, were changed; cf. BCBS (2001a). The problem is that the regulatory rules and the different accounting standards are not fully consistent. Therefore, a bank has to compare the amount of total eligible provisions with the total expected losses amount. If the EL exceeds the provisions, the difference has to be deducted such that it is guaranteed that the total capital amount captures both the UL and the EL; cf. BCBS (2005a), § 43.

[94]BCBS (2005a), § 468.

[95]Cf. BCBS (2004a). Interestingly, the supervisors in the United States proposed a concrete function for mapping the ELGD into the DLGD: $DLGD = 0.08 + 0.92 \cdot ELGD$. Thus, the downturn LGD was a linear mapping from [0%, 100%] to [8%, 100%]. However, in the final rule this supervisory mapping function is not included because of several points of criticism. Nevertheless, the agencies still believe that the formula is an appropriate way to deal with problems in estimating downturn LGDs; cf. FDIC (2007), Sect. III.B.3, p. 69310. However, there is no direct link between this mapping function and the conditional LGD as presented in (2.97).

[96]Cf. Heithecker (2007), p. 31 f., p. 57 ff., and p. 235 ff., for details regarding the maturity adjustment including an outline of the corresponding literature.

to $SF = 1.06$, which is based on the data of the Quantitative Impact Study 3 (QIS 3).[97]

Taking all these points together, the *capital requirement for each credit* under Basel II (in absolute terms) can be expressed as[98]

$$\begin{aligned} UL_{\text{abs},i}^{(\text{Basel})} &= VaR_{\text{abs},i}^{(\text{Basel})} - EL_{\text{abs},i}^{(\text{Basel})} \\ &= EAD_i \cdot \left[DLGD_i \cdot \Phi\left(\frac{\Phi^{-1}(PD_i) + \sqrt{\rho_i} \cdot \Phi^{-1}(0.999)}{\sqrt{1-\rho_i}} \right) - ELGD_i \cdot PD_i \right] \\ &\quad \cdot \frac{1 + (M_i - 2.5) \cdot b}{1 - 1.5 \cdot b} \cdot 1.06 \end{aligned} \tag{2.98}$$

with $b = [0.11852 - 0.05478 \cdot \ln(PD_i)]^2$. Furthermore, the *correlation parameter* is specified by the regulatory framework. Dependent on the asset class (and for some asset classes dependent on the PD and revenue, too), the correlation parameter is between 3% and 24%.[99] For corporate, sovereign, and bank exposures (C,S,B), ρ_i is between 12% (if the PD is very high) and 24% (if the PD is very low):[100]

$$\rho_i^{(\text{C,S,B})} = 0.12 \cdot \frac{1 - \exp(-50 \cdot PD_i)}{1 - \exp(-50)} + 0.24 \cdot \left(1 - \frac{1 - \exp(-50 \cdot PD_i)}{1 - \exp(-50)} \right). \tag{2.99}$$

For small- and medium-sized entities (SMEs), a firm-size adjustment is made. Depending on the total annual sales S_i (in millions of Euros), the correlation parameter will be reduced linearly between 4% (for $S_i \le 5$) and 0% (for $S_i = 50$):[101]

$$\rho_i^{(\text{SME})} = \rho_i^{(\text{C,S,B})} - 0.04 \cdot \left(1 - \frac{\max(S,5) - 5}{45} \right), \tag{2.100}$$

[97]In total, 365 banks participated in the study, which focused on the impact of the Basel II proposals on the minimum capital requirements compared to Basel I; cf. BCBS (2003b).

[98]Cf. BCBS (2005a), § 272, § 273, § 328, § 329, and § 330. The maturity adjustment is only applied to corporate, sovereign, and bank exposures, including small- and medium-sized entities (SMEs). This can also be interpreted as a fixed maturity of $M_i = 1$ year for retail exposures.

[99]For internal purposes a bank could measure ρ from default series or from equity values; cf. Gordy and Heitfield (2002), Düllmann and Trapp (2005), or Lopez (2004). The results for estimating ρ from portfolio data may differ from the correlations given in Basel II, see e.g. Düllmann and Scheule (2003) or Dietsch and Petey (2002), but overall the parameters given in Basel II are reasonable, see especially Lopez (2004).

[100]Cf. BCBS (2005a), § 272. The concrete definition of corporate exposures can be found in BCBS (2005a), § 218 ff.; sovereign and bank exposures are defined in § 229 and § 230.

[101]Cf. BCBS (2005a), § 273.

which leads to a reduction of capital requirements for SMEs. For residential mortgage exposures the correlation is fixed to 15%,[102] for qualifying revolving retail exposures to 4%,[103] and for other retail exposures the correlation parameter is between 3% and 16%:[104]

$$\rho_i^{(\text{Retail})} = 0.03 \cdot \frac{1-\exp(-35 \cdot PD_i)}{1-\exp(-35)} + 0.16 \cdot \left(1 - \frac{1-\exp(-35 \cdot PD_i)}{1-\exp(-35)}\right). \tag{2.101}$$

Taking (2.98) into consideration, the parameters *EAD*, *PD*, *LGD*, and *M* have to be determined. As the complexity of these estimations and the data requirement would be too high for many banks, there exist two versions of the IRB Approach for corporate, sovereign, and bank exposures, as mentioned in Sect. 2.1. In the *Advanced IRB Approach*, all of these parameters have to be estimated by the bank. In the *Foundation IRB Approach*, the LGD and maturity are given by the regulatory rules. Furthermore, only the current outstandings and the commitments have to be determined by the bank, the credit conversion factor and therefore the EAD does not have to be estimated. Thus, under the Foundation Approach, the only parameter that has to be estimated by the bank is the PD.[105] However, for retail exposures, there is no distinction between a Foundation and Advanced IRB Approach. In the *IRB-Retail-Approach*, the parameters EAD, PD, and LGD have to be estimated by the bank.[106] However, in contrast to the IRB Approaches of the other asset classes, in the IRB-Retail-Approach it is allowed to pool credits with similar characteristics such as risk characteristics, collaterals and exposures.[107] As the parameter estimates for the retail portfolio can be based on these risk pools instead of individual borrower grades,[108] the minimum complexity of the IRB-Retail-Approach is significantly lower than of the Advanced IRB Approach.

[102]Cf. BCBS (2005a), § 328.

[103]Cf. BCBS (2005a), § 329.

[104]Cf. BCBS (2005a), § 330.

[105]Cf. BCBS (2005a), § 246 f.

[106]Cf. BCBS (2005a), § 252. The definition of retail exposures can be found in BCBS (2005a), § 231 ff.

[107]Cf. BCBS (2005a), § 401 f.

[108]Cf. BCBS (2005a), § 446.

2.8 Appendix

2.8.1 Alternative Representation of the ES as an Indicator Function

Proposition. *The definition of the ES (2.19) is equal to (2.29):*[109]

$$\frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot 1_{\{\tilde{L}\geq q_\alpha\}}\right] - q_\alpha\left[\mathbb{P}\left[\tilde{L}\geq q_\alpha\right] - (1-\alpha)\right]\right) = \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot 1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\right]\right) \tag{2.102}$$

with

$$1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}} = \begin{cases} 1_{\{\tilde{L}\geq q_\alpha\}} & \text{if } \mathbb{P}\left[\tilde{L}=q_\alpha\right]=0, \\ 1_{\{\tilde{L}\geq q_\alpha\}} - \frac{\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]}\cdot 1_{\{\tilde{L}=q_\alpha\}} & \text{if } \mathbb{P}\left[\tilde{L}=q_\alpha\right]>0. \end{cases} \tag{2.103}$$

Proof. For the case $\mathbb{P}\left[\tilde{L}=q_\alpha\right]=0$, the left-hand side immediately equals the right-hand side of (2.102). Therefore, only the case $\mathbb{P}\left[\tilde{L}=q_\alpha\right]>0$ is analyzed:

$$\begin{aligned} ES_\alpha\left(\tilde{L}\right) &= \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot 1_{\{\tilde{L}\geq q_\alpha\}}\right] - q_\alpha\left[\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)\right]\right) \\ &= \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot 1_{\{\tilde{L}\geq q_\alpha\}}\right] - \tilde{L}\cdot 1_{\{\tilde{L}=q_\alpha\}}\cdot\left(\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)\right)\right) \\ &= \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot 1_{\{\tilde{L}\geq q_\alpha\}} - \frac{\tilde{L}\cdot 1_{\{\tilde{L}=q_\alpha\}}\cdot\left(\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)\right)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]}\right]\right) \\ &= \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot\left(1_{\{\tilde{L}\geq q_\alpha\}} - \frac{\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]}1_{\{\tilde{L}=q_\alpha\}}\right)\right]\right) \\ &= \frac{1}{1-\alpha}\left(\mathbb{E}\left[\tilde{L}\cdot 1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\right]\right), \end{aligned} \tag{2.104}$$

which is the proposed right-hand side of (2.102).

Additionally, we want to show some properties of the function $1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}$, which are useful for analyzing the axioms of coherency. The expected value of this variable is

[109] Cf. Acerbi et al. (2001), p. 8 f.

$$
\begin{aligned}
\mathbb{E}\left[1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\right] &= \mathbb{E}\left[1_{\{\tilde{L}\geq q_\alpha\}} - \frac{\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]}\cdot 1_{\{\tilde{L}=q_\alpha\}}\right] \\
&= \mathbb{E}\left[1_{\{\tilde{L}\geq q_\alpha\}}\right] - \mathbb{E}\left[\frac{\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]}\cdot 1_{\{\tilde{L}=q_\alpha\}}\right] \\
&= \mathbb{P}\left[\tilde{L}\geq q_\alpha\right] - \left(\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)\right) \\
&= 1-\alpha.
\end{aligned} \tag{2.105}
$$

Moreover, we want to show that $1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\in[0,\ 1]$. For $\tilde{L}\neq q_\alpha(\tilde{L})$ this is obvious by the definition of the indicator function. However, for $\tilde{L}=q_\alpha(\tilde{L})$, the variable is given as

$$
\begin{aligned}
1^{\alpha}_{\{\tilde{L}\geq q_\alpha\}}\Big|_{\tilde{L}=q_\alpha} &= 1-\frac{\mathbb{P}\left[\tilde{L}\geq q_\alpha\right]-(1-\alpha)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]} \\
&= 1-\frac{\mathbb{P}\left[\tilde{L}>q_\alpha\right]+\mathbb{P}\left[\tilde{L}=q_\alpha\right]-(1-\alpha)}{\mathbb{P}\left[\tilde{L}=q_\alpha\right]} \\
&= \frac{-\mathbb{P}\left[\tilde{L}>q_\alpha\right]+(1-\alpha)}{\mathbb{P}\left[\tilde{L}\leq q_\alpha\right]-\mathbb{P}\left[\tilde{L}<q_\alpha\right]} \\
&= \frac{\mathbb{P}\left[\tilde{L}\leq q_\alpha\right]-\alpha}{\mathbb{P}\left[\tilde{L}\leq q_\alpha\right]-\mathbb{P}\left[\tilde{L}<q_\alpha\right]}\in[0,\ 1],
\end{aligned} \tag{2.106}
$$

because of $\mathbb{P}\left[\tilde{L}<q_\alpha\right]\leq\alpha\leq\mathbb{P}\left[\tilde{L}\leq q_\alpha\right]$.

2.8.2 Application of Itô's Lemma

An Itô-process is given as

$$dA_t = a(A_t,t)dt + b(A_t,t)dW_t. \tag{2.107}$$

With $a(A_t,t)=\mu\cdot A_t$ and $b(A_t,t)=\sigma\cdot A_t$, we get the stochastic process of the asset value (see (2.54))

$$dA_t = \mu\cdot A_t\,dt + \sigma\cdot A_t\,dW_t. \tag{2.108}$$

Therefore, the asset value follows an Itô-process and Itô's Lemma can be applied in order to determine $dY_t = d\ln A_t$. When Y_t is a function of A_t and t, so we write $Y_t = g(A_t,t)$, Itô's Lemma shows that[110]

[110]Cf. Hull (2006), p. 273 f.

$$\begin{aligned} dY_t &= dg(A_t, t) \\ &= \left(\frac{dg}{dA_t} \cdot a(A_t, t) + \frac{dg}{dt} + \frac{1}{2}\frac{d^2 g}{dA_t{}^2} \cdot b^2(A_t, t)\right) dt + \frac{dg}{dA_t} \cdot b(A_t, t)\, dW_t. \end{aligned} \tag{2.109}$$

With $dY_t = d\ln A_t$, $a(A_t, t) = \mu \cdot A_t$, and $b(A_t, t) = \sigma \cdot A_t$, this leads to

$$\begin{aligned} dY_t &= d\ln A_t \\ &= \left(\frac{1}{A_t} \cdot \mu \cdot A_t + 0 + \frac{1}{2}\left(-\frac{1}{A_t{}^2}\right)(\sigma \cdot A_t)^2\right) dt + \frac{1}{A_t}(\sigma \cdot A_t) dW_t \\ &= \left(\mu - \frac{1}{2}\sigma^2\right) dt + \sigma\, dW_t. \end{aligned} \tag{2.110}$$

2.8.3 Application of Bayes' Theorem for Continuous Distributions

The definition of probability density functions and Bayes' theorem for continuous distributions lead to[111]

$$\begin{aligned} \mathbb{P}(\tilde{y} < u) &= \int_{y=-\infty}^{u} f_Y(y)\, dy \\ &= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{u} f_{X,Y}(x, y)\, dy\, dx \\ &= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{u} f_Y(y|x)\, dy\, f_X(x)\, dx \\ &= \int_{x=-\infty}^{\infty} \mathbb{P}(\tilde{y} < u|x) f_X(x)\, dx. \end{aligned} \tag{2.111}$$

Thus, using $\frac{dF_X(x)}{dx} = f_X(x)$ we get

$$\mathbb{P}(\tilde{y} < u) = \int_{x=-\infty}^{\infty} \mathbb{P}(\tilde{y} < u|x)\, dF_X(x). \tag{2.112}$$

[111]Cf. Tarantola (2005), p. 20.

2.8.4 Limit Distribution and Probability Density Function in the Vasicek Model

In the following, the integral of the distribution (2.74) of the binomial model shall be solved for the limit $n \to \infty$:[112]

$$\begin{aligned} F^{(\infty)}(l) &= \lim_{n\to\infty} F^{(n)}(l) \\ &= \lim_{n\to\infty} \sum_{k=0}^{\lfloor l \cdot n/LGD \rfloor} \int_{x=-\infty}^{\infty} \binom{n}{k} \cdot (p(x))^k \cdot (1-p(x))^{n-k} d\Phi(x) \end{aligned} \tag{2.113}$$

with

$$p(x) = \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot x}{\sqrt{1-\rho}}\right). \tag{2.114}$$

Using $p(x) =: s$ and the identity $\Phi(-y) = 1 - \Phi(y)$, it follows

$$\begin{aligned} s &= \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot x}{\sqrt{1-\rho}}\right) \\ &\Leftrightarrow \sqrt{1-\rho} \cdot \Phi^{-1}(s) = \Phi^{-1}(PD) - \sqrt{\rho} \cdot x \\ &\Leftrightarrow x = -\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho} \cdot \Phi^{-1}(s) - \Phi^{-1}(PD)\right) \\ &\Leftrightarrow \Phi(x) = 1 - \Phi\left(\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho} \cdot \Phi^{-1}(s) - \Phi^{-1}(PD)\right)\right) =: 1 - W(s). \end{aligned} \tag{2.115}$$

Using $d\Phi(x) = d(1 - W(s)) = -dW(s)$ and $\lim_{x\to-\infty} s = \lim_{x\to-\infty} p(x) = 1$ as well as $\lim_{x\to\infty} s = \lim_{x\to\infty} p(x) = 0$, the integral (2.113) can be written as

$$\begin{aligned} F^{(\infty)}(l) &= \lim_{n\to\infty} \int_{s=1}^{0} \sum_{k=0}^{\lfloor l \cdot n/LGD \rfloor} \binom{n}{k} \cdot s^k \cdot (1-s)^{n-k} \cdot (-1) dW(s) \\ &= \int_{s=0}^{1} \lim_{n\to\infty} \sum_{k=0}^{\lfloor l \cdot n/LGD \rfloor} \binom{n}{k} \cdot s^k \cdot (1-s)^{n-k} dW(s). \end{aligned} \tag{2.116}$$

[112]The derivation is based on Vasicek (1991). In contrast to the original paper the derivation is not restrained to the gross loss but includes deterministic $LGD \neq 1$.

The integrand of (2.116) is binomially distributed. According to the central limit theorem of Lindberg-Lévy or the special case for binomial distributions of Moivre–Laplace, this distribution converges to a normal distribution for $n \to \infty$:[113]

$$\begin{aligned}
\lim_{n\to\infty} \sum_{k=0}^{\lfloor l\cdot n/LGD \rfloor} \binom{n}{k} \cdot s^k \cdot (1-s)^{n-k} &= \lim_{n\to\infty} \Phi\left(\frac{n\cdot l/LGD - n\cdot s}{\sqrt{n\cdot s\cdot(1-s)}}\right) \\
&= \lim_{n\to\infty} \Phi\left(\frac{\sqrt{n}}{\sqrt{s\cdot(1-s)}}\left(\frac{l}{LGD} - s\right)\right) \\
&= \begin{cases} \Phi(\infty) = 1 & \text{if } l/LGD > s, \\ \Phi(0) = 1/2 & \text{if } l/LGD = s, \\ \Phi(-\infty) = 0 & \text{if } l/LGD < s. \end{cases}
\end{aligned} \tag{2.117}$$

Therefore, using $W(0) = \Phi(-\infty) = 0$,[114] the distribution (2.116) is equal to

$$\begin{aligned}
F^{(\infty)}(l) &= \int_{s=0}^{1} \lim_{n\to\infty} \sum_{k=0}^{\lfloor l\cdot n/LGD \rfloor} \binom{n}{k} \cdot s^k \cdot (1-s)^{n-k} dW(s) \\
&= \int_{s=0}^{l/LGD} 1\, dW(s) \\
&= W(s)\big|_{s=0}^{l/LGD} \\
&= W\left(\frac{l}{LGD}\right) \\
&= \Phi\left(\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{l}{LGD}\right) - \Phi^{-1}(PD)\right)\right).
\end{aligned} \tag{2.118}$$

The corresponding probability density function $f^{(\infty)}(l)$ is the first derivative of $F^{(\infty)}(l)$. With $d\Phi(y)/dy = \varphi(y)$, $d\Phi^{-1}(y)/dy = 1/\varphi\left(\Phi^{-1}(y)\right)$, and $\varphi(y) = \left(1/\sqrt{2\pi}\right)\cdot\exp(-y^2/2)$ this leads to

[113]See Billingsley (1995), p. 357 f.

[114]Cf. (2.115).

$$
\begin{aligned}
f^{(\infty)}(l) &= \frac{dF^{(\infty)}(l)}{dl} \\
&= \varphi\left(\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{l}{LGD}\right)-\Phi^{-1}(PD)\right)\right)\cdot\sqrt{\frac{1-\rho}{\rho}}\cdot\frac{1}{\varphi\left(\Phi^{-1}\left(\frac{l}{LGD}\right)\right)} \\
&= \sqrt{\frac{1-\rho}{\rho}}\exp\left(-\frac{1}{2\rho}\cdot\left[\sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{l}{LGD}\right)-\Phi^{-1}(PD)\right]^2\right. \\
&\quad \left. +\frac{1}{2}\left[\Phi^{-1}\left(\frac{l}{LGD}\right)\right]^2\right).
\end{aligned}
\tag{2.119}
$$

2.8.5 VaR and ES of the Limit Distribution in the Vasicek Model

According to (2.14), the VaR for continuous distributions can be expressed as

$$VaR_\alpha\left(\tilde{L}\right) = F_L^{-1}(\alpha). \tag{2.120}$$

Thus, corresponding to distribution (2.77), the VaR can be computed as follows:

$$
\begin{aligned}
F^{(\infty)}\left(VaR_\alpha^{(\infty)}\left(\tilde{L}\right)\right) &= \Phi\left(\frac{1}{\sqrt{\rho}}\left(\sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{VaR_\alpha^{(\infty)}\left(\tilde{L}\right)}{LGD}\right)-\Phi^{-1}(PD)\right)\right) \overset{!}{=} \alpha \\
\Leftrightarrow \sqrt{1-\rho}\cdot\Phi^{-1}\left(\frac{VaR_\alpha^{(\infty)}\left(\tilde{L}\right)}{LGD}\right) &= \Phi^{-1}(PD)+\sqrt{\rho}\cdot\Phi^{-1}(\alpha) \\
\Leftrightarrow VaR_\alpha^{(\infty)}\left(\tilde{L}\right) &= \Phi\left(\frac{\Phi^{-1}(PD)+\sqrt{\rho}\cdot\Phi^{-1}(\alpha)}{\sqrt{1-\rho}}\right)\cdot LGD.
\end{aligned}
\tag{2.121}
$$

In order to determine the ES, the representation of (2.20) is used:

$$ES_\alpha\left(\tilde{L}\right) = \frac{1}{1-\alpha}\int_{u=\alpha}^{1} q^u du. \tag{2.122}$$

With (2.121) and using the substitution $y := -\Phi^{-1}(u)$ so that $du/dy = -\varphi(y)$, $y(u=\alpha) = -\Phi^{-1}(\alpha)$ and $y(u=1) = -\Phi^{-1}(1) = -\infty$, this leads to

$$ES_\alpha^{(\infty)}\left(\tilde{L}\right) = \frac{1}{1-\alpha}\int_{u=\alpha}^{1} LGD\cdot\Phi\left(\frac{\Phi^{-1}(PD)+\sqrt{\rho}\cdot\Phi^{-1}(u)}{\sqrt{1-\rho}}\right)du$$

$$
\begin{aligned}
&= \frac{1}{1-\alpha} \cdot LGD \cdot \int_{y=-\Phi^{-1}(\alpha)}^{-\infty} \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot y}{\sqrt{1-\rho}}\right) \cdot (-1) \cdot \varphi(y)dy \\
&= \frac{1}{1-\alpha} \cdot LGD \cdot \int_{y=-\infty}^{-\Phi^{-1}(\alpha)} \Phi\left(\frac{\Phi^{-1}(PD) - \sqrt{\rho} \cdot y}{\sqrt{1-\rho}}\right) \cdot \varphi(y)dy.
\end{aligned}
\tag{2.123}
$$

With the identity[115]

$$
\int_{y=-\infty}^{z} \Phi\left(\frac{x - a \cdot y}{\sqrt{1-a^2}}\right) \cdot \varphi(y)dy = \Phi_2(x, z, a), \tag{2.124}
$$

where $\Phi_2(\cdot)$ is the bivariate cumulative normal distribution as defined in (2.81), (2.123) can be expressed as[116]

$$
ES_\alpha^{(\infty)}\left(\tilde{L}\right) = \frac{1}{1-\alpha} \cdot LGD \cdot \Phi_2\left(\Phi^{-1}(PD), -\Phi^{-1}(\alpha), \sqrt{\rho}\right). \tag{2.125}
$$

2.8.6 Alternative Representation of the Bivariate Normal Distribution

Proposition. *The bivariate normal distribution can be represented as*

$$
\int_{y=-\infty}^{z} \Phi\left(\frac{x - a \cdot y}{\sqrt{1-a^2}}\right) \cdot \varphi(y)dy = \Phi_2(x, z, a). \tag{2.126}
$$

Proof. From

$$
\int_{y=-\infty}^{z} \Phi\left(\frac{x - a \cdot y}{\sqrt{1-a^2}}\right) \cdot \varphi(y)dy = \frac{1}{2\pi} \int_{y=-\infty}^{z} \int_{u=-\infty}^{\frac{x-a\cdot y}{\sqrt{1-a^2}}} \exp\left(-\frac{1}{2}y^2\right) \cdot \exp\left(-\frac{1}{2}u^2\right) du\, dy \tag{2.127}
$$

and using the substitution $u := \frac{w-a\cdot y}{\sqrt{1-a^2}}$ so that $\frac{du}{dw} = \frac{1}{\sqrt{1-a^2}}$, $w(u=-\infty) = -\infty$ and $w\left(u = \frac{x-a\cdot y}{\sqrt{1-a^2}}\right) = x$, we obtain[117]

[115] See Appendix 2.8.6.

[116] See also Pykhtin (2004).

[117] The definition of the bivariate standard normal CDF used in the last step is given in (2.81).

$$\begin{aligned}
&\frac{1}{2\pi}\int\limits_{y=-\infty}^{z}\int\limits_{u=-\infty}^{\frac{x-a\cdot y}{\sqrt{1-a^2}}}\exp\left(-\frac{1}{2}y^2\right)\cdot\exp\left(-\frac{1}{2}u^2\right)du\,dy\\
&=\frac{1}{2\pi}\int\limits_{y=-\infty}^{z}\int\limits_{w=-\infty}^{x}\exp\left(-\frac{1}{2}y^2\right)\cdot\exp\left(-\frac{1}{2}\left(\frac{w-a\cdot y}{\sqrt{1-a^2}}\right)^2\right)\cdot\frac{1}{\sqrt{1-a^2}}dw\,dy\\
&=\frac{1}{2\pi\sqrt{1-a^2}}\int\limits_{y=-\infty}^{z}\int\limits_{w=-\infty}^{x}\exp\left(-\frac{1}{2}\left(y^2+\frac{w^2-2ayw+a^2y^2}{1-a^2}\right)\right)dw\,dy\\
&=\frac{1}{2\pi\sqrt{1-a^2}}\int\limits_{y=-\infty}^{z}\int\limits_{w=-\infty}^{x}\exp\left(-\frac{1}{2(1-a^2)}\left(y^2-2ayw+w^2\right)\right)dw\,dy\\
&=:\Phi_2(x,z,a).
\end{aligned} \tag{2.128}$$

2.8.7 Application of the Strong Law of Large Numbers

Proposition. *The portfolio loss is almost surely equal to the conditional expected loss*

$$\mathbb{P}\left(\lim_{n\to\infty}\left[\tilde{L}-\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right]=0\right)=1 \tag{2.129}$$

under the conditions of infinite granularity (2.83) and (2.84).[118]

Proof. The proof is based upon a version of the strong law of large numbers. For an independent random sequence $\tilde{Z}_i$ the following almost sure convergence holds[119]

$$\mathbb{P}\left(\lim_{n\to\infty}\left[\frac{1}{a_n}\sum_{i=1}^{n}\tilde{Z}_i\right]=0\right)=1 \quad \forall x\in\mathbb{R} \tag{2.130}$$

if

$$\lim_{n\to\infty}a_n=\infty \tag{2.131}$$

[118] The following proof is similar to Gordy (2003), p. 223 f. and Bluhm et al. (2003), p. 88 f.

[119] See Petrov (1996), p. 209, Theorem 6.6.

and

$$\sum_{n=1}^{\infty}\left(\frac{\mathbb{V}(\tilde{Z}_n)}{a_n{}^2}\right) < \infty. \tag{2.132}$$

The random sequence $\tilde{Z}_i$ can be defined as $\tilde{Z}_i := EAD_i \cdot \left(\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} - \mathbb{E}\left[\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}|\tilde{x}\right]\right)$. As it is required that the $\tilde{Z}_i$s are independent, the strong law of large numbers is applied conditional on the realization of the systematic factor $\tilde{x} = x$. Under this condition, the products $\left(\widetilde{LGD}_i \cdot 1_{\{D_i\}}\right)$ are independent by assumption and therefore the $\tilde{Z}_i$'s are independent as well. Defining $a_n := \sum_{j=1}^{n} EAD_j$, the condition (2.131) directly follows from the first granularity assumption (2.83). In order to check the second condition, the boundedness of $\tilde{Z}_n$ is analyzed. The loss variable $1_{\{\tilde{D}_n\}}$ only takes the values one and zero. The LGD is assumed to be in the interval $[-1,\ 1]$.[120] As a consequence, the product $\left(\widetilde{LGD}_n \cdot 1_{\{D_n\}}\right)$ is bounded to $[-1,\ 1]$ and $\left(\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} - \mathbb{E}\left[\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}|\tilde{x}\right]\right)$ is restricted to $[-2,\ 2]$, leading to $\mathbb{V}(\tilde{Z}_n) \leq 4 \cdot EAD_n{}^2$. Therefore, the second condition (2.132) can be written as

$$\sum_{n=1}^{\infty}\left(\frac{\mathbb{V}(\tilde{Z}_n)}{a_n{}^2}\right) \leq \sum_{n=1}^{\infty} 4 \cdot \left(\frac{EAD_n}{\sum\limits_{j=1}^{n} EAD_j}\right)^2 < \infty. \tag{2.133}$$

The last expression is valid due to the second granularity condition (2.84). Thus, the strong law of large numbers (2.130) can be applied. With

$$\begin{aligned}
\frac{1}{a_n}\sum_{i=1}^{n}\tilde{Z}_i &= \frac{1}{\sum\limits_{j=1}^{n} EAD_j}\sum_{i=1}^{n}\left(EAD_i \cdot \left(\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} - \mathbb{E}\left[\widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}|\tilde{x}\right]\right)\right) \\
&= \sum_{i=1}^{n}\left(w_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}} - \mathbb{E}\left[w_i \cdot \widetilde{LGD}_i \cdot 1_{\{\tilde{D}_i\}}|\tilde{x}\right]\right) \\
&= \sum_{i=1}^{n}\left(\tilde{L}_i - \mathbb{E}\left[\tilde{L}_i|\tilde{x}\right]\right) \\
&= \tilde{L} - \mathbb{E}\left[\tilde{L}|\tilde{x}\right]
\end{aligned} \tag{2.134}$$

[120] Negative LGDs are permitted to allow short positions.

this leads to

$$\mathbb{P}\Big(\lim_{n\to\infty}\big(\tilde{L}-\mathbb{E}\big[\tilde{L}|\tilde{x}\big]\big)=0|\tilde{x}=x\Big)=1 \quad \forall x\in\mathbb{R}. \tag{2.135}$$

Using (2.135) it can be shown that the almost sure convergence is also true in the unconditional case:

$$\begin{aligned}\mathbb{P}\Big(\lim_{n\to\infty}\big(\tilde{L}-\mathbb{E}\big[\tilde{L}|\tilde{x}\big]\big)=0\Big) &= \int \mathbb{P}\Big(\lim_{n\to\infty}\big(\tilde{L}-\mathbb{E}\big[\tilde{L}|\tilde{x}\big]\big)=0|\tilde{x}=x\Big)d\mathbb{P}(x)\\ &= \int d\mathbb{P}(x)=1.\end{aligned} \tag{2.136}$$

This completes the proof of (2.129).

2.8.8 Application of Kronecker's Lemma

Proposition. *Assumption (2.83) and (2.84) lead to*

$$\lim_{n\to\infty}\sum_{i=1}^{n} {w_i}^2 = 0. \tag{2.137}$$

Proof. The following proof is based upon Kronecker's Lemma.[121] Let τ_n be a sequence satisfying

$$0<\tau_1\leq\tau_2\leq \quad \text{and} \quad \lim_{n\to\infty}\tau_n=\infty. \tag{2.138}$$

If

$$\sum_{n=1}^{\infty} z_n<\infty, \tag{2.139}$$

then

$$\lim_{n\to\infty}\frac{1}{\tau_n}\sum_{i=1}^{n}\tau_i\cdot z_i=0. \tag{2.140}$$

[121] See Petrov (1996), p. 209, Lemma 6.11.

With $\tau_n := \left(\sum_{j=1}^{n} EAD_j\right)^2$ the conditions (2.138) for τ_n are fulfilled due to the first granularity assumption (2.83). Using $z_n := \left(\frac{EAD_n}{\sum_{j=1}^{n} EAD_j}\right)^2$, (2.139) is valid due to the second granularity assumption (2.84). Therefore, Kronecker's Lemma can be applied, which leads to

$$\begin{aligned}
\lim_{n\to\infty}\frac{1}{\tau_n}\sum_{i=1}^{n}\tau_i\cdot z_i &= \lim_{n\to\infty}\left(\frac{1}{\left(\sum\limits_{j=1}^{n} EAD_j\right)^2}\sum_{i=1}^{n}\left[\left(\sum_{j=1}^{i} EAD_j\right)^2\cdot\left(\frac{EAD_i}{\sum\limits_{j=1}^{i} EAD_j}\right)^2\right]\right)\\
&= \lim_{n\to\infty}\left(\frac{\sum\limits_{i=1}^{n} EAD_i^{\,2}}{\left(\sum\limits_{j=1}^{n} EAD_j\right)^2}\right)\\
&= \lim_{n\to\infty}\left(\sum_{i=1}^{n}\left(\frac{EAD_i}{\sum\limits_{j=1}^{n} EAD_j}\right)^2\right)\\
&= \lim_{n\to\infty}\left(\sum_{i=1}^{n} w_i^{\,2}\right) = 0,
\end{aligned} \tag{2.141}$$

which is (2.137).

2.8.9 Identity of the VaR in the ASRF Model

Proposition. *The following equality is true:*

$$VaR_\alpha\left(\mathbb{E}\left[\tilde{L}|\tilde{x}\right]\right) = \mathbb{E}\left(\tilde{L}|\tilde{x} = VaR_{1-\alpha}(\tilde{x})\right). \tag{2.142}$$

Proof. Using the notation $\mathbb{E}(\tilde{L}|\tilde{x}) =: g\circ\tilde{x}$,[122] with $g(\tilde{x}) = \mathbb{E}(\tilde{L}|\tilde{x})$, and assuming that the conditional expectation is continuously and strictly monotonously decreasing in x, then there exists a unique inverse g^{-1}, which allows the following transformations:[123]

[122] The notation $g\circ\tilde{x}$ means that some function g is composed with $\tilde{x}$.

[123] See Gordy (2003), p. 207 f., for a similar proof.

$$\begin{aligned} & g \circ \tilde{x} \leq g \circ x \\ & \Leftrightarrow g^{-1} \circ g \circ \tilde{x} \geq g^{-1} \circ g \circ x \\ & \Leftrightarrow \tilde{x} \geq x \end{aligned} \tag{2.143}$$

and

$$\inf\{g \circ x\} = g \circ \sup\{x\}. \tag{2.144}$$

Using the definition of the VaR (2.15) this leads to the proposition:

$$\begin{aligned} VaR_\alpha\left(\mathbb{E}\left[\tilde{L}|\tilde{x}\right]\right) &= VaR_\alpha(g \circ \tilde{x}) \\ &= \inf\{l|\mathbb{P}[g \circ \tilde{x} > l] \leq 1 - \alpha\} \\ &= \inf\{g \circ x|\mathbb{P}[g \circ \tilde{x} > g \circ x] \leq 1 - \alpha\} \\ &= \inf\{g \circ x|\mathbb{P}[\tilde{x} < x] \leq 1 - \alpha\} \\ &= g \circ \sup\{x|\mathbb{P}[\tilde{x} < x] \leq 1 - \alpha\} \\ &= g \circ \inf\{x|\mathbb{P}[\tilde{x} > x] \leq 1 - \alpha\} \\ &= g \circ VaR_{1-\alpha}(\tilde{x}) \\ &= \mathbb{E}\left(\tilde{L}|\tilde{x} = VaR_{1-\alpha}(\tilde{x})\right). \end{aligned} \tag{2.145}$$

2.8.10 Identity of the ES in the ASRF Model

Proposition. *For $n \to \infty$, the ES of the portfolio loss converges to the ES of the conditional expected loss:*

$$\lim_{n\to\infty} ES_\alpha\left(\tilde{L}\right) = ES_\alpha\left(\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right). \tag{2.146}$$

Proof. If it is assumed that the loss distribution is continuous, the second term of ES definition (2.19) vanishes.[124] Therefore it only has to be shown that

$$\lim_{n\to\infty} \mathbb{E}\left[\tilde{L} \cdot 1_{\left\{\tilde{L} \geq q_\alpha\left(\tilde{L}\right)\right\}}\right] - \mathbb{E}\left[\mathbb{E}\left(\tilde{L}|\tilde{x}\right) \cdot 1_{\mathbb{E}\left(\tilde{L}|\tilde{x}\right) \geq q_\alpha\left(\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right)}\right] = 0. \tag{2.147}$$

With $\tilde{X} := \tilde{L} - q_\alpha\left(\tilde{L}\right)$ the first term can be written as

[124] Gordy (2003) shows that it is no necessary condition that the loss distribution has to be continuous. If some additional properties, especially regarding the continuity of the conditional expected loss and of the distribution of the systematic factor, are fulfilled in an interval of x that contains $VaR_\alpha(\tilde{x})$, it follows that $\lim_{n\to\infty} \mathbb{P}\left(\tilde{L} \geq q_\alpha\right) = 1 - \alpha$ so that the second term of the ES definition still vanishes. See Gordy (2003), p. 228 f.

$$\begin{aligned}\mathbb{E}\left[\tilde{L}\cdot 1_{\left\{\tilde{L}\geq q_\alpha\left(\tilde{L}\right)\right\}}\right] &= \mathbb{E}\left[\left(\tilde{L}-q_\alpha\left(\tilde{L}\right)\right)\cdot 1_{\left\{\tilde{L}\geq q_\alpha\left(\tilde{L}\right)\right\}}\right]+q_\alpha\left(\tilde{L}\right)\cdot\mathbb{E}\left[1_{\left\{\tilde{L}\geq q_\alpha\left(\tilde{L}\right)\right\}}\right]\\ &= \mathbb{E}\left[\max\left(\tilde{X},0\right)\right]+q_\alpha\left(\tilde{L}\right)\cdot\mathbb{P}\left[\tilde{L}\geq VaR_\alpha\left(\tilde{L}\right)\right].\end{aligned} \tag{2.148}$$

Using the shorter notation $\mu(\tilde{x}) := \mathbb{E}(\tilde{L}|\tilde{x})$ and with $\tilde{Y} := \mu(\tilde{x}) - \mu(q_{1-\alpha}(\tilde{x}))$ as well as $\mu(q_{1-\alpha}(\tilde{x})) = q_\alpha(\mu(\tilde{x}))$ from (2.90), the second term of (2.147) equals

$$\begin{aligned}&\mathbb{E}\left[\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\cdot 1_{\left\{\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\geq q_\alpha\left(\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right)\right\}}\right]\\ &\quad = \mathbb{E}\left[\mu(\tilde{x})\cdot 1_{\{\mu(\tilde{x})\geq q_\alpha(\mu(\tilde{x}))\}}\right]\\ &\quad = \mathbb{E}\left[(\mu(\tilde{x})-\mu(q_{1-\alpha}(\tilde{x})))\cdot 1_{\{\mu(\tilde{x})\geq\mu(q_{1-\alpha}(\tilde{x}))\}}\right]+\mu(q_{1-\alpha}(\tilde{x}))\cdot\mathbb{E}\left[1_{\{\mu(\tilde{x})\geq\mu(q_{1-\alpha}(\tilde{x}))\}}\right]\\ &\quad = \mathbb{E}\left[\max\left(\tilde{Y},0\right)\right]+\mu(q_{1-\alpha}(\tilde{x}))\cdot\mathbb{P}[\mu(\tilde{x})\geq\mu(q_{1-\alpha}(\tilde{x}))].\end{aligned} \tag{2.149}$$

Thus, (2.147) can be written as

$$\begin{aligned}&\lim_{n\to\infty}\mathbb{E}\left[\tilde{L}\cdot 1_{\left\{\tilde{L}\geq q_\alpha\left(\tilde{L}\right)\right\}}\right]-\mathbb{E}\left[\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\cdot 1_{\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\geq q_\alpha\left(\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right)}\right]\\ &\quad = \lim_{n\to\infty}\left(\mathbb{E}\left[\max\left(\tilde{X},0\right)\right]+q_\alpha\left(\tilde{L}\right)\cdot\mathbb{P}\left[\tilde{L}\geq VaR_\alpha\left(\tilde{L}\right)\right]\right)\\ &\qquad -\left(\mathbb{E}\left[\max\left(\tilde{Y},0\right)\right]+\mu(q_{1-\alpha}(\tilde{x}))\cdot\mathbb{P}[\mu(\tilde{x})\geq\mu(q_{1-\alpha}(\tilde{x}))]\right)\\ &\quad = \lim_{n\to\infty}\left(\mathbb{E}\left[\max\left(\tilde{X},0\right)-\max\left(\tilde{Y},0\right)\right]\right.\\ &\qquad \left.+q_\alpha\left(\tilde{L}\right)\cdot\mathbb{P}\left[\tilde{L}\geq VaR_\alpha\left(\tilde{L}\right)\right]-\mu(q_{1-\alpha}(\tilde{x}))\cdot\mathbb{P}[\mu(\tilde{x})\geq\mu(q_{1-\alpha}(\tilde{x}))]\right).\end{aligned} \tag{2.150}$$

Using

$$\lim_{n\to\infty} q_\alpha\left(\tilde{L}\right) = \mu(q_{1-\alpha}(\tilde{x})) \tag{2.151}$$

from (2.91) and[125]

$$\lim_{n\to\infty}\mathbb{P}\left[\tilde{L}\geq q_\alpha\left(\tilde{L}\right)\right] = \mathbb{P}\left[\mu\left(\tilde{L}\right)\geq\mu\left(q_\alpha\left(\tilde{L}\right)\right)\right] = 1-\alpha, \tag{2.152}$$

the last two terms of (2.150) vanish:

[125] Cf. footnote 118.

$$\begin{aligned}
&\lim_{n\to\infty}\left(q_\alpha\left(\tilde{L}\right)\cdot\mathbb{P}\left[\tilde{L}\geq VaR_\alpha\left(\tilde{L}\right)\right]-\mu(q_{1-\alpha}(\tilde{x}))\cdot\mathbb{P}[\mu(\tilde{x})\geq\mu(q_{1-\alpha}(\tilde{x}))]\right)\\
&\quad=\lim_{n\to\infty}\left[q_\alpha\left(\tilde{L}\right)-\mu(q_{1-\alpha}(\tilde{x}))\right]\cdot(1-\alpha)\\
&\quad=0. \qquad (2.153)
\end{aligned}$$

Additionally, the inequality $-|x-y|\leq\max(x,0)-\max(y,0)\leq|x-y|$ holds $\forall x,y\in\mathbb{R}$. Using this inequality and (2.151), the remaining first term of (2.150) can be evaluated:

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{E}\left[\max\left(\tilde{X},0\right)-\max\left(\tilde{Y},0\right)\right]&\leq\lim_{n\to\infty}\left|\mathbb{E}\left[\tilde{X}-\tilde{Y}\right]\right|\\
&=\lim_{n\to\infty}\left|\mathbb{E}\left[\tilde{L}-q_\alpha\left(\tilde{L}\right)-[\mu(\tilde{x})-\mu(q_{1-\alpha}(\tilde{x}))]\right]\right|\\
&=\lim_{n\to\infty}\left|\mathbb{E}\left[\tilde{L}-\mu(\tilde{x})\right]-\left[q_\alpha\left(\tilde{L}\right)-\mu(q_{1-\alpha}(\tilde{x}))\right]\right|\\
&=\lim_{n\to\infty}\left|\mathbb{E}\left(\tilde{L}\right)-\mathbb{E}\left(\mathbb{E}\left(\tilde{L}|\tilde{x}\right)\right)-0\right|\\
&=0 \qquad (2.154)
\end{aligned}$$

and

$$\lim_{n\to\infty}\mathbb{E}\left[\max\left(\tilde{X},0\right)-\max\left(\tilde{Y},0\right)\right]\geq-\lim_{n\to\infty}\left|\mathbb{E}\left[\tilde{X}-\tilde{Y}\right]\right|=0. \qquad (2.155)$$

Thus, the first term vanishes, too, which completes the proof of (2.146).

http://www.springer.com/978-3-7908-2606-7

Risk Management in Credit Portfolios
Concentration Risk and Basel II
Hibbeln, M.
2010, XX, 248 p., Hardcover
ISBN: 978-3-7908-2606-7
A product of Physica-Verlag Heidelberg