DANMARKS NATIONALBANK

DANMARKS NATIONALBANK WORKING PAPERS 2005 • 27

Lisbeth Borup
Financial Markets
Danmarks Nationalbank, Copenhagen

and

Dorte Kurek
Financial Markets
Danmarks Nationalbank, Copenhagen

and

Anne Dyrberg Rommer
Financial Markets
Danmarks Nationalbank, Copenhagen
and
Centre for Applied Microeconometrics (CAM),
Institute of Economics, University of Copenhagen

Assessing the consequences of Basel II: Are there incentives for cherry-picking when banks pool data across countries?

21 June 2005

Please direct any enquiries to
Danmarks Nationalbank, Information Desk, Havnegade 5, DK-1093 Copenhagen K
Denmark
Tel.: +45 33 63 70 00 (direct) or +45 33 63 63 63
Fax : +45 33 63 71 03
E-mail:info@nationalbanken.dk

Henvendelser kan rettes til :
Danmarks Nationalbank, Informationssektionen, Havnegade 5, 1093 København K.
Telefon: 33 63 70 00 (direkte) eller 33 63 63 63
E-mail: info@nationalbanken.dk

ISSN (trykt/print) 1602-1185

ISSN (online) 1602-1193

Resumé

Baselkomiteens reviderede anbefalinger til kapitaldækningsregler (Basel II) træder i kraft i 2007. Basel II giver bankerne mulighed for at benytte deres interne modeller til at estimere sandsynligheden for misligholdelse (probability of default, PD) ved beregning af minimumskapitalkravet ved hjælp af de interne rating-metoder. Valide estimater af PD kræver en betydelig mængde data og misligholdelsesobservationer. Basel II tillader, at banker pooler deres data med henblik på at få et tilstrækkeligt datagrundlag. Det har ført til en række internationale data pooling-projekter. Således behøver selv internationale banker mere data for at opfylde Basel II's krav.

Så vidt vi ved, har ingen papirer indtil nu sammenlignet bankernes kapitalkrav beregnet på basis af PD estimeret via kreditrisikomodeller estimeret for et enkelt land (enkelt-lands model) med kreditrisikomodeller estimeret med data fra flere lande (fler-lande model) samt diskuteret de incitamentsstrukturer estimation af de forskellige modeller kan give for banker, som pooler data.

Formålet med dette papir er at illustrere konsekvenserne for de beregnede kapitalkrav af at poole data fra flere lande ved estimation af PD. Vi konstruerer en hypotetisk portefølje af lån til små og mellemstore virksomheder for en hypotetisk bank med forretning i Frankrig, Italien og Spanien. Hertil benytter vi virkelige data udtrukket fra den pan-europæiske Amadeus-database fra Bureau van Dijk. På basis af disse data beregnes PD i kreditrisikomodeller estimeret på de enkelte lande samt på basis af kreditrisikomodeller estimeret med poolede data fra de tre lande (fler-lande modeller). De estimerede PD'er bruges herefter til at beregne minimumskapitalkravet.

Resultaterne viser, at bankerne kan have incitament til at vælge en bestemt metode, hvis de pooler data, fordi den resulterer i et lavere kapitalkrav. Det beregnede kapitalkrav for den hypotetiske bank varierer med op til 18 pct. afhængig af den valgte metode, mens kapitalkravet opgjort på de enkelte lande varierer op til 47 pct. Resultaterne har især interesse for banker med forretning i flere lande, som planlægger at poole data fra deres udlån i landene med henblik på at estimere PD, fx fordi der ikke er tilstrækkeligt data for det enkelte land. De er ligeledes interessante for banker, der planlægger at poole data med banker fra andre lande, hvis de ikke selv har tilstrækkeligt med data.

Selv om vores definition af misligholdelse er ens for de tre lande, og vi har kontrolleret for variable såsom alder, størrelse, juridisk form og sektor for hver virksomhed, finder vi relativt store forskelle i de resulterende minimumskapitalkrav for porteføljen i hvert af de tre lande, når PD estimeres ved hjælp af enkelt-lande modeller og fler-lande modeller. Vi viser, at det ikke er nok for bankerne at benytte samme definition af misligholdelse og samme regnskabsprincipper i landene. Når banker pooler data, skal banker og tilsynsmyndigheder også undersøge modellerne, herunder hvilke faktorer der forårsager misligholdelse.

Abstract

The Basel Committee's Revised Framework for Capital Measurement and Capital Standards (Basel II) will enter into force in 2007. Basel II facilitates the use of banks' internal models to estimate probability of default (PD) when calculating the minimum capital requirement using the internal ratings-based approaches. Valid estimates of the PDs require a considerable amount of data and default observations. Basel II allows for banks to pool their data to overcome their data shortcomings and a number of international data pooling projects have emerged. Thus even international banks need more data to fulfil the requirements of Basel II.

To the best of our knowledge, so far no study has compared the banks' capital requirements calculated on the basis of PDs estimated from single-country credit-scoring models and multi-country credit-scoring models and accordingly no study has discussed the incentive structure this might create for banks pooling data.

The purpose of this paper is to illustrate the consequences on the calculated capital requirements of pooling data for estimation of PD from several countries. We construct a hypothetical portfolio of loans to small and medium sized enterprises for a hypothetical bank operating in France, Italy and Spain. For this purpose we use real world data extracted from the pan-European Amadeus database provided by Bureau van Dijk. Using this data, the PDs are estimated on the basis of single-country credit-scoring models and on the basis of multi-country credit-scoring models with pooled data from the three countries. The estimated PDs are then used to calculate the minimum capital requirements.

The result shows that there might be incentives for cherry-picking, i.e. that banks are motivated to choose a certain method because it results in a lower capital requirement. The calculated capital requirements vary with up to 18 percent depending on the choice of method for the hypothetical bank. Calculated for the individual countries it varies up to 47 percent. The results are of particular interest for banks operating in several countries, which plan to pool data from the various countries in order to estimate PDs, maybe due to lack of a sufficient single-country database. They are equally interesting for banks planning to pool data with banks from other countries to make up for an insufficient database.

Though our default definition is the same for the three countries and we have controlled for variables such as age, size, legal form and sector of each firm, we find quite large differences in terms of the resulting minimum capital requirements for the portfolio in each of the three countries, when the PDs are estimated using a single-country credit-scoring model compared to using multi-country credit-scoring models. We show that it is not enough for banks to apply similar definitions of default and similar accounting regimes in the countries. Banks and regulators should also have a careful look into the models, especially the factors that drive financial distress, when banks pool data.

Lisbeth Borup[1]
Dorte Kurek[2]
Anne Dyrberg Rommer[3]

Assessing the consequences of Basel II: Are there incentives for cherry-picking when banks pool data across countries?[4]

[1] Lisbeth Borup, Financial Markets, Danmarks Nationalbank, Havnegade 5, DK-1093 Copenhagen, Denmark. Phone: +45 3363 6363. Email: lb@nationalbanken.dk.

[2] Dorte Kurek, Financial Markets, Danmarks Nationalbank, Havnegade 5, DK-1093 Copenhagen, Denmark. Phone: +45 3363 6363. Email: dku@nationalbanken.dk.

[3] Anne Dyrberg Rommer (corresponding author), Financial Markets, Danmarks Nationalbank, Havnegade 5, DK-1093 Copenhagen, Denmark, phone: +45 3363 6363, email: ady@nationalbanken.dk, and Centre for Applied Microeconometrics (CAM), Institute of Economics, University of Copenhagen, email: anne.dyrberg@econ.ku.dk.

[4] The authors would like to thank Hans Christian Kongsted, Jesper Berg, Karsten Biltoft, Anders Møller Christensen, Michael Friis, Hugo Frey Jensen, Kristian Kjeldsen, Jens Lundager and Jakob W. Lund for commenting on earlier versions of the paper.

1. Introduction

With the new capital adequacy rules, Basel II, entering into force in 2007, banks worldwide including EU banks will be given the opportunity to apply their internal models (credit-scoring models) when calculating their minimum capital requirement for credit risk using the internal ratings-based approaches (IRB). A bank's use of internal models has to be approved by the supervisor and the bank must demonstrate to the supervisor, that it fulfills a number of requirements, including validation requirements. Valid estimates and "backtesting" of credit-scoring models require a considerable amount of data and default observations. Many banks are still in the early phase of building up the necessary database in order to fulfill the model requirements of Basel II. Basel II allows for banks to pool their data to overcome their data shortcomings, c.f. BCBS, Basel Committee for Banking Supervision, (2004:86ff). Specific recommendations on the setting up and use of pools are not given, e.g. the use of cross-country pools for banks operating in several countries or for banks wishing to pool data with similar banks in other countries.

Following the Basel Committee's work on Basel II a number of data pooling projects have emerged illustrating that many banks require more data to fulfill the IRB requirements of Basel II. To name a few projects, a group of European banks incl. Barclays Capital, Calyon, Royal Bank of Scotland, JP Morgan Chase, and NIB Capital has formed the Pan-European Credit Data Consortium and plan to share loss data for their commercial loan portfolio, c.f. Dunbar (2005). Furthermore, Standard&Poor's is coordinating the pooling of loss data on project finance for Citigroup, ABN Amro, Société Générale and Deutsche Bank, c.f. Cass (2002).

The purpose of this paper is to illustrate the consequences on the calculated capital requirements of pooling data from several countries for banks' estimation of probability of default (PD), when following the foundation IRB approach in Basel II (further details on the foundation IRB approach are found in section 3.1 and in appendix 1). We construct a hypothetical loan portfolio for a hypothetical bank operating in France, Italy and Spain. For this purpose we use real world data on French, Italian and Spanish small and medium-sized enterprises (SMEs) extracted from the Amadeus database provided by Bureau van Dijk.[5] The Amadeus database is arguably the best available database for cross-country analysis of firms in financial distress and comprises harmonized accounting data and financial distress events. Using this data, the PDs are estimated on the basis of single-country credit-scoring models and on the basis of multi-country credit-scoring models with pooled data from the three countries (with and without country dummies). The estimated PDs are then used to calculate the minimum capital requirements under the foundation IRB approach in Basel II. The consequences of the two setups (single-country versus multi-country credit-scoring models) are discussed.

The results are of particular interest for banks operating in different countries, which plan to pool data from their exposures in the various countries in order to estimate PDs like our hypothetical bank, maybe due to lack of a sufficient single-country database. The results are equally interesting for banks planning to pool data with banks from other countries to estimate PDs to make up for an insufficient database. In

[5] The Amadeus database is based on public information. The information in the data base differs from the data, which is available to individual banks, e.g. individual banks can use a 90 days past due default definition and include more parameters in their models than we are able to.

this respect it should be highlighted that Italy, Spain and France are countries, which in important aspects are fairly alike. They all belong to Continental Europe and they are all members of the European Monetary Union. Furthermore, they are inspired by the same legal tradition, c.f. La Porta, Lopez-De-Silanes, Shleifer and Vishny (1998).

The harmonized dataset presented and analyzed in Rommer (2005a) is used here. The consequences of estimating multi-country PDs based on pooled data compared to calculating single-country PDs in terms of the resulting capital requirements are investigated for French, Italian and Spanish small and medium-sized enterprises. To the best of our knowledge, this is the only study, which compares the calculated capital requirements based on PDs from single-country credit-scoring models and multi-country credit-scoring models. Other studies analyze the treatment of SME loans under the Basel II framework in one country only. Fabi, Laviola and Reedtz (2004) provide an empirical evaluation of the impact of Basel II on Italian corporates, Saurina and Trucharte (2003) analyze the impact of Basel II on lending to Spanish small and medium-sized enterprises and Masschelein (2003) analyzes the implication for Belgian banks of the Basel II treatment of SME loans. Thus, this paper fills a gap in the literature.

This paper is structured the following way. First, the role of capital for banks is reviewed from a theoretical point of view. Secondly, the Basel II framework and the internal ratings-based approaches are reviewed, and the minimum requirements for estimation probability of default and guidelines on pooling data are compared to the method we apply. Thirdly, the dataset is described and the estimation of the credit-scoring models is discussed. Fourthly, the capital requirements are calculated using credit-scoring models and multi-country credit-scoring models, and the results are discussed. The last section concludes.

2 The role of capital for banks and the reasoning for capital requirements

This section presents the theoretical background for having capital requirements, and more specifically, the determinants of the capital structure for banks and the reasoning for financial regulators to apply capital adequacy rules.

In contrast to non-financial firms banks are subject to capital adequacy rules set by financial regulators. Thereby financial regulators implicitly assume that the optimal amount of capital for banks' shareholders without any regulations, often referred to as economic capital, is too low.[6]

The capital structure of banks[7] is determined in part by the same variables that determine the capital structures of non-financial firms – taxes, expected costs of financial distress, transaction costs, signaling behavior and agency problems arising

---

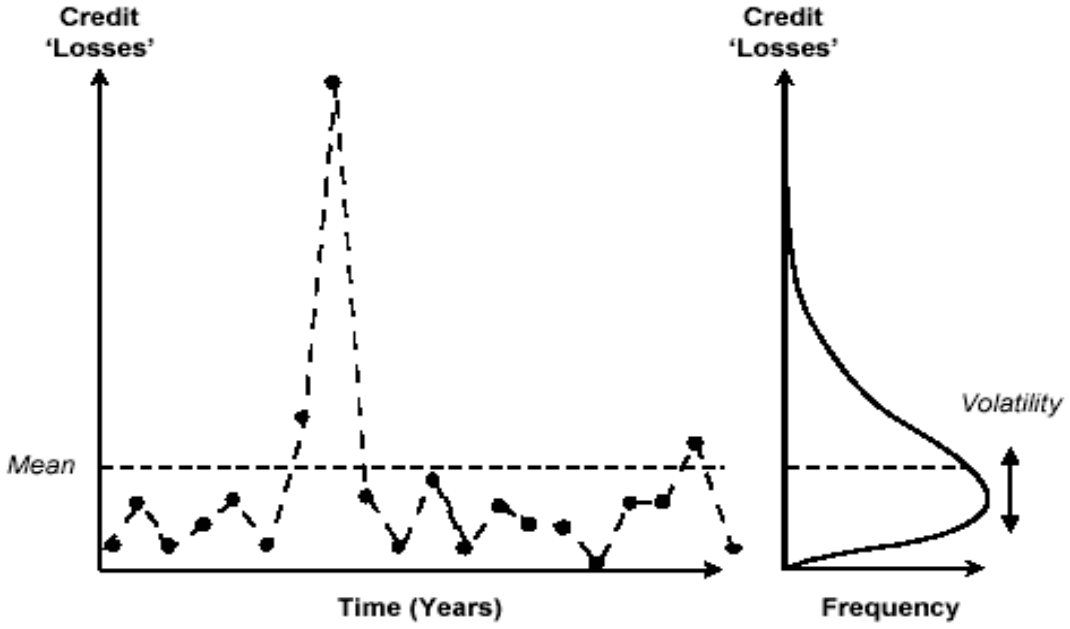
[6] This is one argument. However, one could also argue that banks benefit from the introduction of regulation. E.g. this was (generally) the case for Danish banks: "When the banks were not subject to any regulation – before the first Danish banking act in 1919 – capital adequacy was generally far higher than today. Around the mid-19th century capital and reserves were approximately 40 per cent of the balance-sheet total, and by around 1900 a good 20 per cent. In the 1920s this had fallen to around 12 per cent. Today, the banks' net capital is an average of approximately 6 per cent of the balance-sheet total" (Andersen (2004)).

[7] Berger, Herring and Szegö (1995) provide a comprehensive overview of the theoretical arguments to explain the optimal amount of capital for banks. Kjeldsen (2004) explains why capital requirements for banks are necessary and why banks usually prefer to hold excess reserves.

from asymmetric information between shareholders and creditors and between owners and managers. If raising capital quickly is costly for any of these reasons, then firms may hold additional capital as financial slack to take advantage of unexpected profitable opportunities or to guard against unexpected losses. Banks differ substantially from non-financial firms because they are protected by a regulatory safety net. The existence of deposit insurance causes the depositors to demand no risk premium and thus makes it possible for banks to borrow at the risk free rate no matter the leverage (capital-to-liabilities ratio). Deposit insurance is therefore likely to move the optimal capital structure towards a low level of capital, as debt financing is cheaper for banks than for non-financial firms.

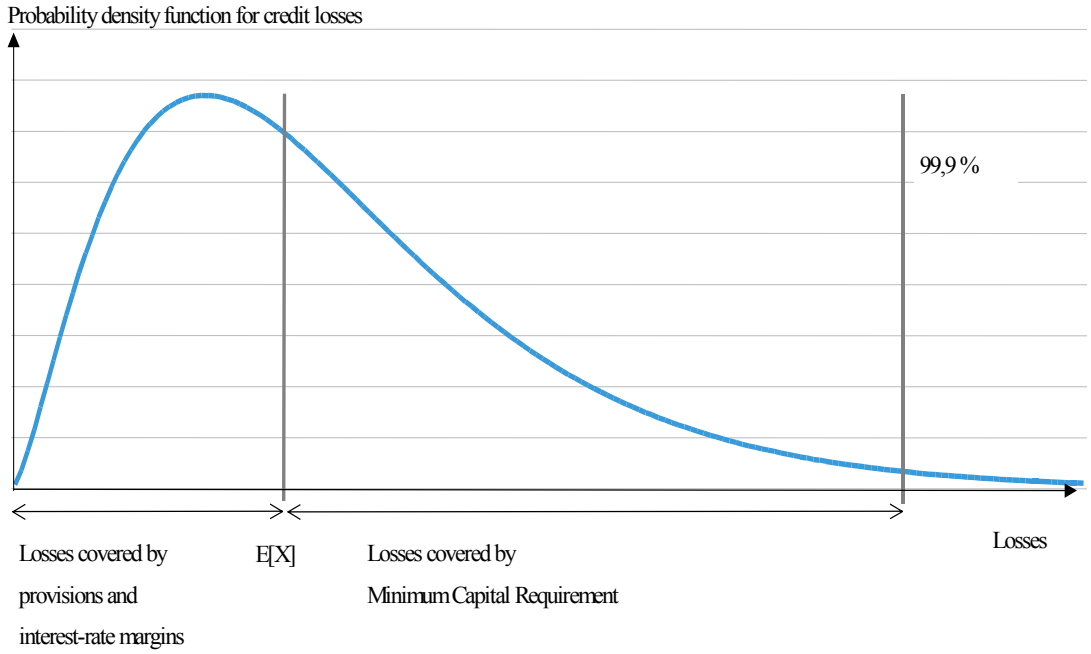
The main instrument for regulating banks is capital requirements, which should ensure that banks have sufficient capital. Regulatory capital requirements are motivated by two main concerns. First, as a means to protect the economy from negative externalities caused by financial problems in one bank spreading to other banks. This could be the case if a bank failure brings on a general distrust in the banking system causing difficulties for banks to raise capital on financial markets and/or causing depositors in a panic to withdraw all their deposits from the banking system (bank runs). Second, as a means to avoid the value of a failed bank's assets dropping below the value of the depositors' claims on the bank. This helps ensuring that the bank can be reconstructed or wound up more easily.

Chart 2.a: The distribution of credit losses



Source: Mercer Oliver Wyman (2005)

Chart 2.b: Probability density function for credit losses



Source: Thoraval and Duchateau (2003) and own manufacture

The concept of credit loss is illustrated in chart 2.a. The expected loss is the mean of all the credit losses of the portfolio, c.f. the chart to the left. The resulting distribution function over the credit losses is shown to the right. The unexpected loss depends on the volatility of credit losses. The distribution of credit losses is characterized by a long tail, which is explained by a relatively high probability of small losses and a small probability of very high losses. Therefore the mean (expected loss) is not located at the maximum of the distribution.

The bank makes provisions and sets its interest-rate margin at a level corresponding to the expected value of the losses, stated as E[X] in chart 2.b. If the losses rise above the mean, provisions and interest-rate margins will not be sufficient, c.f. chart 2.b. The purpose of capital requirements is to ensure that, with a given degree of probability, e.g. 99.9 per cent, the bank's capital should cover the unexpected losses above the mean. That is the likelihood that a loss will exceed the bank's capital is 0.1 per cent. This builds of course on the assumption that the probability distribution can be determined with sufficient accuracy.

Banks have the inherent characteristic of a relative low capital-to-liabilities ratio (high gearing). To encourage prudent management of the risks associated with this unique balance sheet structure, regulatory authorities have from early on introduced certain capital adequacy requirements. When the Basel Committee took the lead in the late 1980s to develop a risk-based capital adequacy standard (Basel I, c.f. BCBS (1988)) the objectives were to strengthen the soundness and stability of the international banking system and, by ensuring a high degree of consistency in the framework's application, to diminish the sources of competitive inequalities among international banks, c.f. van Greuning and Bratanovic (2003). While the new framework of Basel II

aims to provide a comprehensive approach to measuring banking risks, its fundamental objectives remain the same as those of the 1988 Accord.

3. Basel II

This section reviews the Basel II framework and describes the calculation of the minimum capital requirements using the internal ratings-based approaches. Furthermore, the Basel II minimum requirements for estimation of probability of default as well as the Basel II requirements and other guidelines on data pooling are discussed. Finally, we compare our approach to the requirements and guidelines.

3.1 IRB approach of the Basel II framework

In June 2004 the central bank governors and the heads of banking supervisory authorities of the G10 countries (the Basel Committee of Banking Supervision) endorsed the Revised Framework for Capital Measurement and Capital Standards, also known as Basel II, which is a set of recommendations for the capital requirements imposed on banking organizations by supervisory authorities. Basel II is designed to cope with the shortcomings of the current regime, the 1988 Capital Accord, c.f. Caruana (2004a) and Caruana (2004b).

The 1988 Capital Accord states that banks should hold capital in excess of 8 per cent of the risk weighted assets for credit risk and market risk. The Basel Committee of Banking Supervision introduced the Accord as a set of capital adequacy rules to apply for major internationally active banks based in the G10 countries. A common set of rules was necessary to prevent bank failures and at the same time ensure level playing field for banks competing in the same countries. Following the introduction of the Accord more than 100 countries including the EU chose to adopt the 1988 Accord for their banks.

The Basel Committee has made it clear, that Basel II aims at the same overall global capital requirement as the 1988 Accord, but to make each individual bank's capital requirement more closely linked to its risk of economic loss.

More than 100 countries worldwide are expected to adopt Basel II, c.f. BIS (2004) and Keefe (2004). The European Commission has worked in parallel with the Basel Committee on proposals for directives to replace the 1988 Accord with Basel II in the EU countries for credit institutions and investment firms. On 14 July 2004 the European Commission presented its proposals for Directives to transpose the Basel II into European Law. The proposed directives are expected to be finally adopted by the European Parliament and the Council in 2005 and enter into force at end-2006. Credit institutions and investment firms can apply the existing capital-adequacy rules until end-2007. However, institutions applying for the most advanced approaches[8] for calculation of minimum capital requirements may not apply the new rules until 2008.

Basel II consists of three pillars. Pillar I sets out criteria for banking organizations' calculation of minimum capital requirements to cover market risk, credit risk and operational risk. The latter was not covered by the 1988 Accord. Pillar 1 represents an extension of the requirements in the 1988 Accord and introduces more sophisticated calculation approaches, which aligns the minimum capital requirements more closely

[8] The advanced internal ratings-based approach for credit risk and advanced measurement approach for operational risk

to the banks' risk of economic loss, especially for credit risk. For credit risk the banks can with the approval of the supervisory authority choose one of three approaches, namely the standard approach, the foundation internal ratings-based approach and the advanced internal ratings-based approach. In the two sophisticated approaches, the internal ratings-based approaches (IRB), the banks use their internal credit-scoring models in the calculation of the capital requirement, c.f. below. Pillar II requires banks to assess their need for capital in relation to their overall risk profile including risks not or only partly covered by pillar I, e.g. interest-rate risk on the banking book, business risk and strategic risk. Furthermore, the supervisory authority must evaluate the banks' assessment of its capital need. Pillar III sets out principles for banks' disclosure of information concerning risks and capital to enhance market discipline.

The focus of this paper is on the foundation IRB-approach for calculation of the capital requirement for credit risk of exposures to small and medium sized enterprises (SME) under pillar I. Nonetheless we will briefly review the main features of both the foundation IRB approach and the advanced IRB approach.

Both approaches use four quantitative inputs: 1) the probability of default (PD), which measures the likelihood that a borrower will default over a one-year time horizon, 2) loss given default (LGD), which measures the proportion of the exposure that will be lost if a default occurs, 3) exposure at default (EAD), which measures the nominal value of the debt and 4) the effective maturity (M), which measures the remaining economic maturity of exposure. Risk weights are calculated by inserting PD, LGD and M into the formulas prepared by the Basel Committee. The minimum capital requirement (K^*) for each exposure can then be calculated as 8 % of the risk weight (RW) multiplied by EAD:

$$K^* = 0.08 \times RW \times EAD$$

The main difference between the foundation and the advanced IRB approach is the extent the approaches rely on inputs provided by credit institutions on the basis of their own estimates, as opposed to those inputs that are pre-specified by the supervisor. In the foundation IRB approach, only the PDs are estimated by the credit institutions, whereas, in the advanced IRB approach, the credit institutions estimate all four risk factors themselves.

Under the IRB approaches, banks must categorize credit exposures into the following broad classes: corporate, sovereign, bank, retail and equity. For corporate exposures banks are permitted to distinguish separately exposures to SMEs and give them a lower risk weight. SMEs are defined by their size as corporate exposures where the reported sales for the consolidated group of which the firm is a part is less than 50 million euro. A firm-size adjustment is made to the corporate risk weight formula for exposures to SMEs. Very small exposures to SMEs can under certain conditions be eligible for the more favourable retail treatment. For an SME to be treated as a retail exposure it needs to have its loans managed as other retail exposures and the total exposure of a bank to an individual firm has to be less than 1 million euro. For the interested reader the Basel II formulas for calculation of credit exposures to SMEs are presented in appendix 1.

In this paper the calculations of minimum capital requirements for credit risk for SMEs are based on estimates of PD's from a credit-scoring model. We apply the foundation IRB approach which prescribes LGD=45 % and M=2.5 years for corporate

exposures. EAD is calculated as the sum of loans and long-term debt[9]. We assume that exposures are without collateral. SMEs are defined as corporate exposures where the reported sales of the firms are between 1 and 50 million euro. In this paper we do not consider the more favorable retail treatment for the very small SMEs with exposures less than 1 million euro. That is, we apply the SME formula for all firms in the data set.

3.2 Basel II requirements for estimation of probability of default

For supervisory authorities to approve a bank for one of the IRB approaches for calculation of the minimum capital requirement for credit risk, the bank must demonstrate that it fulfills a number of requirements on an ongoing basis. This section describes the minimum requirements in Basel II, which are of particular importance when estimating probability of defaults, and it compares these requirements to the data and method we apply for calculating minimum capital requirements.

According to Basel II default has occurred when 1) the bank considers that the obligor is unlikely to pay its credit obligations to the bank, the parent undertaking or any of its subsidiaries in full without recourse by the bank to actions such as realizing security, or 2) the obligor is past due more than 90 days on any material credit obligation to the bank, the parent undertaking or any of its subsidiaries. Basel II lists a number of indicators for the bank's assessment of an obligor's unlikeliness to pay. These include situations where the bank puts the credit obligation on non-accrued status, makes a value adjustment resulting from a significant perceived decline in credit quality, sells the credit obligation at a material credit-related economic loss, consents to a distressed restructuring of the credit obligation likely to result in diminished financial obligation, has filed for obligor's bankruptcy or where the obligor has sought or has been placed in bankruptcy. The same definition is used in the EU directive proposal.

Of the two complementing approaches for defining default, the suggested indicators for the banks' assessment of the obligor's unlikeliness to pay (approach 1) imply that in most cases default will occur before the obligor is 90 days past due (approach 2).

According to the EU directive proposal until 2012 the 90-days may be extended up to 180 days if local conditions make it appropriate, c.f. European Commission (2004: article 154, 4).

The bank must have a rating system of its obligors with a meaningful distribution of exposures across grades with minimum 7 rating grades for non-defaulted obligors and one for defaulted obligors. For each grade the bank must estimate a PD for all the obligors in that grade. The EU directive proposal relaxes this requirement and allows banks to use direct estimates of PDs for calculation of the capital requirement, i.e. without introducing a rating system, c.f. European Commission (2004: Annex VII, part 4, 4). This implies that each obligor has a separate PD, whereas, when using a rating system, the obligors of each rating grade have the same PD.

According to Basel II the PD should be a long run average of one-year default rates. The length of the underlying historical period should be at least five years. In the EU

---

[9] According to Basel II, with the foundation internal ratings-based approach, EAD should be calculated as the on and off-balance sheet position gross of specific provisions or partial write-offs, c.f. BCBS (2004:66).

directive proposal, the requirement for the underlying historical period is cut down to two years until end-2007, increasing by one year per year thereafter until end-2010, c.f. European Commission (2004:article 154,4). Therefore, the end-requirement will be an underlying period of five years as in Basel II.

Basel II requires banks to use information and techniques that take appropriate account of the long-run experience when estimating the average PD for each rating grade. The EU directive proposal allows banks to use direct estimates of PDs without introducing a rating system. Nevertheless, we expect that these estimates should also take appropriate account of the long-run experience.

The number of exposures in the sample and the data period used for quantification has to be sufficient to provide the bank with confidence in the accuracy and robustness of its estimates. In order to avoid over-optimism, a bank must add to its PD-estimates a margin of conservatism. If methods and data are less satisfactory and the likely range of errors is larger, the bank has to use a larger margin of conservatism.

Finally, Basel II emphasizes that human judgment and human oversight is necessary to ensure that all relevant information, including that which is outside the scope of the model is also taken into consideration, and that the model is used appropriately. In addition, banks must recognize the importance of judgmental considerations in combining results of techniques and in making adjustments for limitations of techniques and information.

For the purpose of this paper we estimate the probability of default for firms using a credit-scoring model. In line with the EU directive proposal, we estimate probability of default directly for each firm without developing a rating system and assigning firms to risk grades. Introducing a rating system would be more in line with Basel II, but would on the other hand force us to make a number of assumptions (e.g. definition of rating grades and calculation of average probability of defaults) bringing in unnecessary noise, which would blur our results. Furthermore, we apply the model's estimation of credit scores directly as probability of defaults. This implies no use of human judgment concerning probability of default for each firm or adding a conservative margin to avoid over-optimism. Use of human judgment and adding a conservative margin is not relevant for our problem since this would imply a shift in the PDs in the same direction for PDs obtained both from the estimation of the single-country credit-scoring models and from the multi-country credit-scoring models.

The estimations of the credit-scoring models in this paper cover only firms that have handed in financial statements in the period 2000 – 2002. It is important to note that the correct computation of minimum capital requirements, according to Basel II, demands that the underlying historical period must be at least five years. The estimated probability of default in this paper does not take this requirement into account. However, it does accommodate the requirements in the EU directive proposal, in which the underlying historical period is reduced to two years and increasing one year per year from end-2007 to end-2010. This means that in a transition phase it is sufficient to use the data available to us.

The focus in this paper is on the firms that go bankrupt. This "financial distress-event" is a fairly late credit event compared to the Basel II definition. As this model is based on public information only, it is not possible to follow the Basel II default definition. Several studies imply that this is not of significant importance, when building the credit-scoring model. Hayden (2003) shows that credit-scoring models that rely on bankruptcy as default criterion instead of delay-in-payments can be equally powerful in predicting the credit loss events. Furthermore, Moody's Investors Service (2001) reports that experience shows that the factors that can predict default are generally the same, no matter whether the definition of default is 90 days past due or bankruptcy. It is not uncommon to use a fairly late credit event in academic studies. The definition of default used in two of the three studies mentioned in the introduction is the bankruptcy event (Saurina and Trucharte (2003) and Masschelein (2003)). The last study uses banks' classification of loans as bad loans (Fabi, Laviola and Reedtz (2004)).

Though using the Basel II default definition is not essential for building credit-scoring models in general, it is of importance when the probability of defaults estimated in the credit-scoring model is used for calculating the level of the capital requirement – since it affects the level of the PD and hence the resulting capital requirement. For the purpose of this paper, which compares the capital requirements using single-country credit-scoring models and multi-country credit-scoring models and the incentive structures this might create, the absolute level of the calculated capital requirements is not important.

3.3 Data pooling – requirements and guidelines

Basel II and the EU directive proposal both allow banks to pool data to overcome their data shortcomings. The emergence of a number of data pooling projects illustrates the considerable need for banks to pool data in order to fulfill the Basel II IRB requirements. This section describes Basel II's and the EU directive proposal's requirements with regard to data pooling and furthermore it highlights recommendations on data pooling from guidelines published by supervisory authorities. Finally, these requirements and guidelines are compared to the data and method we apply for calculating minimum capital requirements.

Basel II allows for banks to pool internal data with external data, c.f. BCBS (2004:92ff). The bank has to demonstrate that the internal rating systems of other banks in the pool are comparable to its own and representative of the population of the bank's actual borrowers. Furthermore, estimates based on internal or external data should be representative of long-run experience. The Basel Committee does not give specific recommendations on the setting up and use of pools, e.g. the use of cross-country pools - for instance for a bank with exposures in different countries (like our hypothetical bank) or for a bank with data for one country who plans to pool this with banks who have data from other countries.

The EU directive proposal elaborates on the requirements for using external and pooled data. The directive proposal states that credit institutions using external data that is not itself consistent with the definition of default shall demonstrate that appropriate adjustments have been made to achieve broad equivalence with the definition of default, c.f. the European Commission (2004, Annex 7, part 4, 46). In addition, a credit institution using data pooled across credit institutions has to

demonstrate that the pool is representative for the portfolio for which the pooled data is used and that the pooled data is used consistently over time by the credit institution for its permanent estimates, c.f. the European Commission (2004, Annex 7, part 4, 57). The directive proposal requires that credit institutions use internal data for assigning exposures to rating grades as the primary source of information when estimating PDs and LGDs. Credit institutions are permitted to use external data (including pooled data) for quantification provided a strong link can be demonstrated between 1) the credit institution's process for assigning exposures to grades and the process used by the external data source and 2) the credit institution's internal risk profile and composition of the external data, c.f. the European Commission (2004, Annex 7, part 4, 69). That is, the pooled data should be representative for the credit institution's loan portfolio. The directive proposal does not provide rules on the use of external data or pooled data when using direct PD-estimation, i.e. estimating PD's without using rating grades.

The wording of the directive proposal with regard to the use of external data and data pooling can be interpreted as applying for data pooling between banking institutions, which are a part of the same banking group, as well as for one bank wishing to pool data with other banks. The rules would thus apply for our hypothetical bank with cross-country exposures and for a bank planning to pool data with banks in other countries.

Published guidance on the use of pooled data for PD-estimation is very limited or kept in general terms. Oesterreichische Nationalbank (2004:63ff) highlights the importance of a uniform definition of default in the pooled data and points out, that discrepancies can arise between individual countries due to use of different accounting standards.

In this paper we use a uniform definition of default. The default observations are constructed as firms exiting the Amadeus database due to bankruptcy. As the three countries (France, Italy and Spain) are countries with French-civil-law tradition, differences between the institutional frameworks are limited, c.f. La Porta, Lopez-De-Silanes, Shleifer and Vishny (1998), who scores the countries based on enforcement variables (e.g. efficiency of judicial system), accounting standards and creditor rights (e.g. no automatic stay on assets and secured creditors paid first). Furthermore, the Amadeus database harmonizes accounting data from different countries. Discrepancies due to the use of different accounting standards are thus very limited.

The UK Financial Services Authority (UK FSA) states in their consultation paper that if a bank uses data pooled across institutions it should be able to demonstrate to the FSA, that the pool is representative for the portfolio for which the pooled data is used, c.f. UK FSA (2005, appendix 1). This statement could apply both for a bank operating in other countries through subsidiaries, which plans to pool data from the subsidiaries and the parents, and for a bank planning to pool data with banks based in other countries. The consultation paper does not elaborate on the definition of representativity.

The EU Committee of European Banking Supervisors is also working on guidelines on the implementation, validation and review of the IRB approaches including guidelines on the use of data pooling. These guidelines are not yet published.

Our hypothetical bank pools data on SME exposures in France, Italy and Spain. Based on this data set (portfolio) we estimate statistical default prediction models, namely three different hazard models, c.f. section 5. As we control for a number of effects in the estimations, e.g. industry, size, age and legal form (see section 5), we do not need to ensure that our sample from France is representative of the sample from Spain and Italy etc. In fact, even if there were no firms in the manufacturing sector in France, it would not matter for the estimation of the probability of default, as the dummy, which indicates whether a specific company is a manufacturing company or not, would then be set to 0, when the PDs for France are estimated. In the same way, it would not matter if there, for example, were no public limited liability companies in Spain. In the actual estimations, the estimated PD for each individual firm includes information on a wide number of individual characteristics, and so all of these characteristics are taken into consideration, when the credit-scoring models are estimated.

Even though it is not necessary to show that all industries, sizes and legal forms are present for each country in our sample in order to get consistent estimates, c.f. above, for the interested reader, table 3.3 shows that in each of the countries we are analyzing the same industries, sizes and legal forms are present (further details on data are given in section 4). There are differences between the number of firms in the different industries, sizes and legal forms in the respective countries, but as is explained above, it is not a problem in our setting.

Table 3.3: The analyzed sample split up on legal form, size and industries (percentages in brackets)

		Spain	France	Italy
Legal form (number of firms)	Public limited liability company	46,317 (62 pct.)	89,314 (86 pct.)	31,312 (32 pct.)
	Private limited liability company	28,635 (38 pct.)	14,530 (14 pct.)	65,129 (68 pct.)
Size (measured as ln(total assets))	Mean	8.51	8.45	8.61
	Median	8.35	8.30	8.50
Industries (number of firms)	Farming, forestry and fishing	1,616 (2 pct.)	1,284 (1 pct.)	443 (0 pct.)
	Mining	905 (1 pct.)	858 (1 pct.)	722 (1 pct.)
	Manufacturing	26,648 (36 pct.)	31,829 (31 pct.)	50,861 (53 pct.)
	Energy	389 (1 pct.)	275 (0 pct.)	285 (0 pct.)
	Construction	8,945 (12 pct.)	8,808 (8 pct.)	7,298 (8 pct.)
	Trade and hotel	21,554 (29 pct.)	35,714 (34 pct.)	25,215 (26 pct.)
	Transport	4,578 (6 pct.)	6,113 (6 pct.)	3,789 (4 pct.)
	Business service	7,176 (10 pct.)	14,244 (14 pct.)	4,958 (5 pct.)
	Public service activities	1,022 (1 pct.)	2,380 (2 pct.)	1,233 (1 pct.)
	Organisations	2,119 (3 pct.)	2,339 (2 pct.)	1,317 (1 pct.)

In the case where the estimated model does not include a large number of explanatory variables, e.g. only a few accounting ratios, the user of the model would need to ensure that the same sectors are represented in the various portfolios, which are pooled, just as well as it would be a good idea to ensure, that the composition of the portfolio with respect to legal form, would be the same across portfolios. Concerning both variables (sector affiliation and legal form), a number of credit-scoring studies have documented that the probability of default differs across sectors, as well as across legal form, see e.g. Dyrberg (2004).

4. Data

The data used for Italy, Spain and France comes from the Amadeus database, which is a pan-European database provided by Bureau van Dijk. This section presents the data set and explains the construction of the dependent variable. Furthermore, this section gives an overview of the sample selection criteria and it presents the hypothetical loan portfolio. For further details on the data set the reader is referred to Rommer (2005a).

4.1 The raw data

The Amadeus database comprises information on financial issues as well as non-financial issues. Bureau van Dijk has harmonized the database so that the financial items across countries are comparable. As part of the non-financial information, the database entails a legal status variable. This variable contains information on the status of the firm (active, bankrupt etc.). This piece of information is particularly important for this study, as it is used to construct the dependent variable.

Unfortunately, information on the legal status variable is only kept in the database for 3 years, and so, currently, the estimations of the credit-scoring models cover only firms that have handed in financial statements in the period 2000 – 2002. Ideally, the estimation period would have covered a full business cycle.

In the dataset, firms that hand in a financial statement in 2000 are recorded as belonging to year 2000. Firms that hand in a financial statement in 2001 are recorded as belonging to year 2001 etc. Some firms are represented with one data point, e.g. in 2000, in 2001 or in 2002, other firms will be represented by two points, e.g. in 2000 and 2001 or in 2001 and 2002, and other firms will be represented by three data points. In technical terms, the firms are both flow and stock sampled and the length of the spells varies across firms. There is one spell for each firm. When a firm has left the sample, it can never re-enter, i.e. the exit event is an absorbing state. For further details on duration data, the reader is referred to Dyrberg (2004).

4.2 Construction of the dependent variable

The legal status variable for constructing the dependent variable, which is the event "financially distressed firms", i.e. a measure of the firms that may inflict a loss on the financial sector. In the Amadeus database it is registered whether or not the company is bankrupt (France, Italy, Spain), whether or not the company is in receivership (France), and whether or not it has defaulted on its payments (Spain, France). The broadest measure of financial distress, which can be used here, is therefore a measure, which, for each country, includes the events that are registered for the respective countries. This broad measure is not a satisfactory measure for financial distress in this set up, where the impact on the calculated capital requirements of the estimation of single-country credit-scoring models and multi-country credit-scoring models is at

focus. Therefore, in order to make the financial distress event consistent across countries, we include only bankrupt firms for the measure of financial distress in the analyzed hypothetical loan portfolio. Accordingly, a hypothetical loan portfolio is constructed, which only includes the firms that go bankrupt and active firms.

4.3 Sample selection criteria and the hypothetical loan portfolio

In order to construct the hypothetical loan portfolio, various sample selection criteria, which are discussed in details in Rommer (2005a), are applied to the data. Table 4.a gives an overview of the applied criteria. In particular, note that the analysed sample only includes SMEs with annual sales less than 50 million euro to comply with the criteria for when a firm can be treated as SME using the IRB approach.

After the application of the sample selection criteria, the hypothetical loan portfolio is constructed. Table 4.b shows the hypothetical loan portfolio, i.e. the number of observations in the sample, which are used in the estimations (with bankruptcy as a default criterion). From the table we can see that our particular hypothetical bank has experienced most bankruptcies in its French portfolio (597 bankruptcies are registered) and the smallest number of bankruptcies in Spain (115 bankruptcies are registered).

Table 4.a: Sample selection criteria

Criteria	
Conceptual	Only unconsolidated statements are analysed
	Financial institutions and non-financial holding companies are excluded
	Only public limited liabilities and private limited liabilities are analysed
	Only SMEs with at least 10 employees and with total assets of at least 2 million euro. This criterion ensures that micro-companies, which resemble households, are excluded from the sample, and furthermore, that only "truly" active companies are considered, c.f. the discussions in Rommer (2005a) SMEs with total annual sales less than 50 million euro
	Some firms leave with no explanation (that is, they are not assigned an exit code). These firms are called attritioners. Based on the analysis in Rommer (2005a), they are excluded from the dataset
Other	Active companies are excluded if they hand in a financial statement in 2000 and 2002 only.
	Various corrections are made to the database (e.g. firms with illogical variables, such as short-term debt less than zero and a solvency ratio larger than 100 pct., are excluded).
	Firms with missing variables on any of the explanatory variables are excluded.
	If a company hands in two financial statements in one year, only the last financial statement is included in the estimations.

Table 4.b: Bankrupt firms and other firms (period covered 2000-2002)

	Spain		France		Italy		Pooled	
	Number of firm-years	In percent of total	Number of firm-years	In percent of total	Number of firm-years	In percent of total	Number of firm-years	In percent of total
Bankruptcy	115	0.15	597	0.57	155	0.16	867	0.32
Active and censored firms	74837	99.85	103247	99.43	96286	99.84	274370	99.68
Total	74952	100	103844	100	96441	100	275237	100

5. Estimation of the PDs

This section gives an overview of how the PDs are estimated, including the explanatory variables that are used in the estimations.

Based on the dataset presented in table 4.b, accounting-based credit-scoring models are estimated. An accounting-based credit-scoring model is based on information extracted from company accounts and in some cases also non-financial information (such as the age of the company). It estimates the probability that a particular firm will default on its debt obligations. Various estimation techniques have been suggested in the accounting-based credit-scoring literature (e.g. discriminant analysis and logistic regression).[10] Here the estimation strategy of Shumway (2001) is followed, thus the credit-scoring models are estimated as hazard models. The hazard functions are specified as logit models. The firms that exit for other reasons than financial distress (i.e. firms that are voluntarily liquidated) are treated as censored or no longer observed when they leave the dataset. Three different credit-scoring models are estimated: First, individual credit-scoring models for each country are estimated. Second, a multi-country credit-scoring model with country dummies is estimated. Third, a multi-country credit-scoring model without country dummies is estimated.

The explanatory variables, which are included in the estimations, can be seen from table 5. They are divided into three categories: Core variables, proxies and controls. For further details the reader is referred to Rommer (2005a), which uses the same data.

[10] For an overview of the literature the reader is referred to Jones (1987), Dimitras, Zanakis and Zopounidis (1996), Altman and Saunders (1998), Balcaen and Ooghe (2004) and Lando (2004). Some of the often-quoted accountings-based credit-scoring studies are Beaver (1966), Altman (1968), Ohlson (1980) and Shumway (2001).

Table 5: The explanatory variables

	Variable
Core variables	Profitability: Earnings ratio= EBITDA/total assets. EBITDA = earnings before interest, taxes, depreciation and amortization
	Solvency: Solvency = Equity/total assets
	Leverage: Loans/total assets
	Firms size: Ln(total assets)
	Age: The year of the financial statement minus the year of incorporation
Proxies	Legal form: This dummy is equal to 1, if it is a private limited liability company, and equal to 0, if it is a public limited liability company.
	Independence indicator: Three dummies are included. One is equal to one when the ownership is very concentrated (when at least one of the shareholders has an ownership above 49.9 pct.). One is equal to one when the ownership is of medium concentration (when none of the shareholders have an ownership percentage above 49.9 pct, but at least one or more shareholders has an ownership percentage above 24.9 pct.). One is equal to one when the ownership is not so concentrated (when none of the shareholders has more than 24.9 pct. of ownership share). Reference category is all other firms (for which there is no information on the shareholders).
	Shareholders: This variable measures the number of recorded shareholders.
	Subsidiaries: This variable measures the number of subsidiaries that a company has registered.
Controls	Sector affiliation dummies: The data is divided in the following industries: 1) Farming, Forestry, Fishing 2) Mining, 3) Manufacturing, 4) Energy, 5) Construction, 6) Trade and hotel, 7) Transport, 8) Business service, 9) Public service activities, 10) Organisations etc. Financial firms and holding companies are excluded from the analysis. As there are no NACE codes for the IT and tele-sector a (self-constructed) IT and tele-dummy is included in the estimations. On top of belonging to one of the above sectors a firm is considered to by in the IT and tele group if it has activities in one of the sectors listed in table 14.b in section 14 in Rommer (2005a). Further details on the sectors are found in table 14.a in section 14 in Rommer (2005a). Note that in the actual estimations the following industries are grouped (as the data are too sparse otherwise): Organisations and public service activities are grouped. Farming, forestry, fishing, mining and energy are grouped.
	Macroeconomic environment: Year dummies are included to control for business cycle effects. The reference year is 2000. Two dummies are included. One is equal to 1 for the firms, which hand in their financial statements in 2001. One is equal to 1 for the firms, which hand in their financial statements in 2002.
Controls (only in the multi-country credit-scoring model with country dummies)	Country dummies: A dummy for each country is included in the estimations to control for country-specific effects.

Source: Rommer (2005a) and own manufacture

6. Results

Based on data on SMEs in Italy, Spain and France we have constructed a hypothetical loan portfolio. We have applied single-country credit-scoring models and multi-country credit-scoring models with and without country dummies to estimate PDs. The PDs have been used to calculate the resulting minimum capital requirement for our hypothetical bank using the foundation IRB approach.

Table 6: A comparison of the capital requirements for the hypothetical bank using the companies that were active in 2002 in the three model set ups, in million euros, and the largest range in capital requirements in percent of the smallest capital requirement (right column)

	Single-country models	Multi-country model (with country dummies)	Multi-country model (without country dummies)	Largest range in capital requirements in percent of the smallest capital requirement
IT	94,703	99,323	139,654	47
ES	122,588	109,837	148,737	35
FR	100,034	104,489	81,883	27
Total	317,324	313,649	370,274	18

Note: The number of active companies in 2002 is 35,818 in Italy, 29,447 in Spain and 41,251 in France. The total number of active companies in the three samples is 106,516.

This section presents the resulting capital requirements calculated on the basis of the PDs estimated in the single-country credit-scoring models and the multi-country credit-scoring models with and without country dummies, c.f. appendix 2, table 6 and figure 6. The estimated probabilities of default are set into the formulas for calculating the capital requirements using the foundation IRB approach for SMEs. In the formulas loss given default (LGD) is set to 45 pct. and maturity (M) is taken as 2.5 as prescribed by the foundation IRB approach. The exposure at default (EAD) is calculated as the sum of loans and long-term debt.[11] [12]

The results presented in table 6 (further details are found in appendix 2) show that for each country the difference in the calculated capital requirements is quite large, when the probability of default is estimated using a single-country credit-scoring model compared to using multi-country credit-scoring models (with or without country dummies). Particular different results are obtained from the single-country credit-scoring models and the multi-country credit-scoring model without country dummies, and between the multi-country credit-scoring model with country dummies and the multi-country credit-scoring model without dummies, c.f. table 6. It is remarkable how limited the differences in the calculated capital requirements are between the single-country and the multi-country credit-scoring models with country dummies.

It is particularly noticeable that the lowest capital requirement in the countries is obtained using different model set-ups. In the Italian case, estimating a single-country credit-scoring model delivers the lowest capital requirement, in the Spanish case a multi-country model with country dummies delivers the lowest capital requirement and in the French case the multi-country model without country dummies delivers the lowest capital requirement. In the Italian and Spanish case the multi-country model without country dummies delivers the highest capital requirement, whereas the multi-

[11] In the Amadeus database liabilities are spit up on current liabilities (short-term) and non-current liabilities (long-term). Current liabilities are divided into loans, creditors and other current liabilities. Non-current liabilities are divided into long-term debt and other non-current liabilities (incl. provisioning).

[12] A firm can have several bank connections. It is not indicated in the Amadeus database, whether the firms use one or more than one bank connection. In the calculations total exposure of a firm is interpreted as the exposure to one hypothetical bank.

country model with country dummies delivers the highest capital requirements in the French case.

In the situation where a bank considers pooling data with banks from other countries, the bank has an incentive to choose the method, which delivers the lowest capital requirement without considering what level of capital is actually appropriate to cover the overall credit risk, i.e. cherry-picking. Note, however, that Basel II does not allow banks to change their model for estimation of PD every so often for instance in order to obtain the lowest possible minimum capital requirement. Basel II states that banks must monitor the model stability, c.f. BCBS (2004:86). The EU directive proposal furthermore requires banks to validate the accuracy and consistency of rating systems, processes and the estimation of all relevant parameters, and points out that changes in estimation methods and data shall be documented, c.f. European Commission (2004:appendix VII, part 4, 109 and 112).

Table 6 shows that a bank with exposures to Italian firms would benefit form choosing a single-country model, a bank with the exposures to Spanish firms would benefit from using the multi-country model with country dummies etc. Furthermore, table 6 illustrates that our hypothetical bank would obtain the lowest capital requirement from estimating a multi-country model with country dummies. The overall capital requirement would then be marginally lower (313,649 million euro) compared to estimating single-country credit-scoring models (317,324 million euro) and much lower compared to a multi-country credit-scoring model without country dummies (370,274 million euro).

Chart 6: Capital requirements over total exposure at default (EAD), split up on country and method

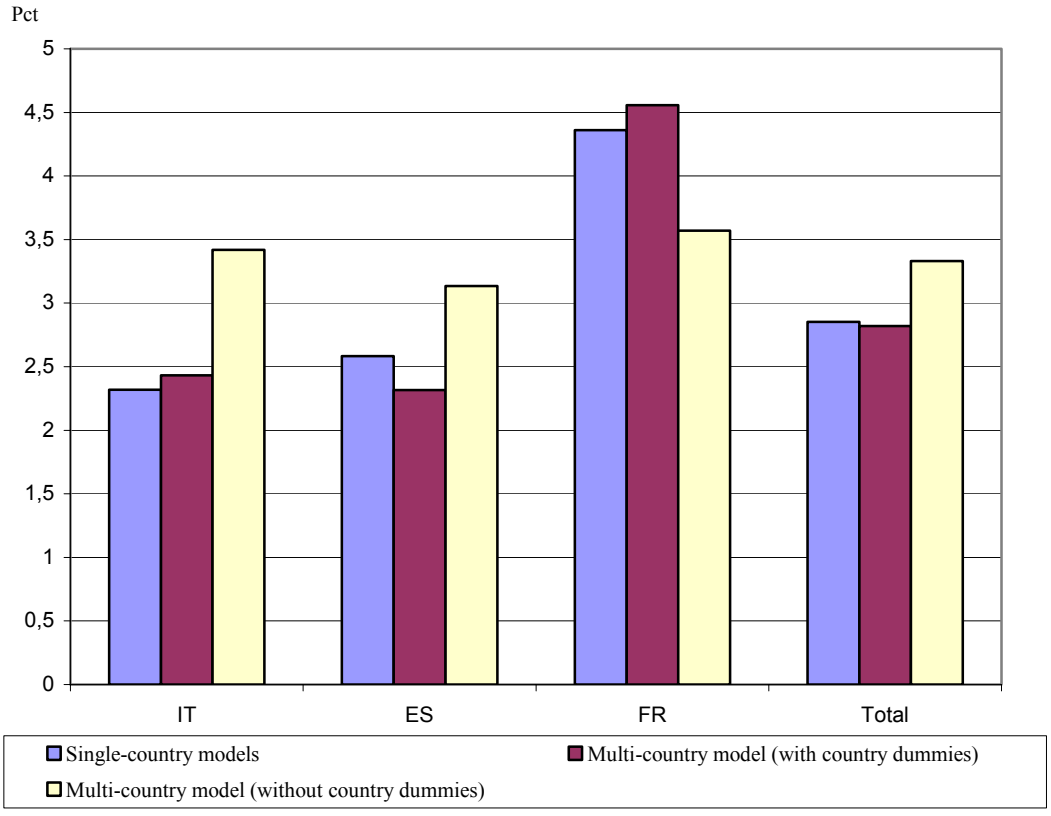


Chart 6 illustrates the capital requirement over total exposure at default (EAD), split up on country and method. As we normalize the capital requirements in each country with the total exposure at default in each country, the figure enables us to compare the riskiness of the exposures across countries. The chart shows that the French loan portfolio is more risky (in the sense that it has a higher capital requirements over EAD), and that the Italian and Spanish loan portfolios are at the same risk-level, when the two preferred methods are compared. As we would expect, the total loan portfolio, which consists of exposures to all countries, is placed somewhere in between the French, and the Italian and Spanish loan portfolios.

The differences in the calculated capital requirements are not due to differences in the default definition, as only bankrupt firms from countries with French-civil-law tradition were considered. However, it can be sensitive towards the way we constructed the default measure. As was noted in section 4.2, the broadest measure of financial distress, which could be used, is a measure, which, for each country, includes the events that are registered for the respective countries. In order to make the default definition comparable across countries, we chose, in this paper, only to focus on bankruptcies. It is also worth pointing out that we used a database, which was harmonized across countries, and that we controlled for a number of variables in the estimations, including age of the company, size, solvency ratio, leverage, profitability, legal form, ownership variables and sector affiliation.

The calculations in this paper are based purely on the quantitative and technical requirements of the foundation IRB approach of Basel II and the EU directive proposal. We do not take into consideration effects of applying human judgment and a conservative perspective on PD estimation. The example of our hypothetical bank serves to illustrate the purely technical consequences of pooling data and the incentives it might give for banks when calculating minimum capital requirements.

One reason why the calculated capital requirements are different in the countries depending on whether the single-country or a multi-country model (with or without country dummies) is estimated could be that the predictors of financial distress differ across countries, c.f. box 6. As we controlled for macroeconomic effects by the use of year dummies, c.f. section 4, the differences in the predictors of financial distress in the single-country models are not due to differences in the macroeconomic environment in the respective countries, i.e. different levels of the real interest rate, growth, inflation etc. Note, however that differences in the macroeconomic environments could have the implication that the number of firms in financial distress differ across countries.

The concrete implication for banking supervision from the analysis is that banks and supervisory authorities must be aware that the pooling of data from several countries should be done with caution. As there are not many official guidelines from authorities on the issue, we believe that the illustration in this paper of the consequences of pooling data from several countries serves as an important input to the debate on how to set up credit-scoring models in banks that have cross-border exposures as well as for banks who choose to pool their data with banks in other countries. We have shown that it is not enough for banks to apply similar definitions of default and similar accounting regimes in the countries. Banks and regulators should also have a careful look into the models, especially the factors that drive financial distress, e.g. along the lines of Rommer (2005a). Rommer (2005a) is one concrete example of an econometric study that investigates the determinants of

financial distress in several countries. However, it is not only important to assess the factors that drive financial distress. Credit institutions and regulators should also pay special attention to 1) the sample selection and design, 2) the statistical technique and 3) the evaluation of results.

Under item 1) one issue, which is important to assess, is the extent of drop-outs in the credit institutions portfolio. Our analysis is based on a panel data set. In our observation window, which spans from 2000 to 2002, we follow the firms from the time, when they are incorporated, till they leave the sample. This is not always the case for individual credit institutions, c.f. Rommer (2005b). Credit institutions may experience drop-outs for a number of reasons, e.g. a firm may choose another bank as it offers a better service or a better price, or e.g. because the specific firm is asked to leave its current bank, as it suspects that it is heading into financial distress. It is important that the drop-outs are carefully analyzed in order to find out what kind of drop-outs one deals with. Otherwise inconsistent estimates may be obtained, when the credit-scoring model is set up and estimated. Another issue, which also falls under item 1), is the reject inference problem: In Basel II it is stated that internal ratings and default and loss estimates must play an essential role in the credit approval process. It is important to be aware, that if the "models estimated using data on already approved applicants are applied to all applicants, then a sample selection bias is introduced", c.f. Rommer (2005b). The problem is that if only obligors, who have already been approved for a loan, are taken into account then it is not appropriate to use the same model to consider new applications. In the academic literature this problem is called the reject inference problem.

Concerning item 2), which is the choice of statistical technique, a wide range of papers discuss the differences between the various techniques, which have been suggested in the literature. The standard credit-scoring methods are multivariate discriminant analysis, logistic regression and hazard models. These methods are discussed in Altman (1968), Ohlson (1980) and Shumway (2001), respectively. It is important that credit institutions and regulators are aware of the advantages and limitations of the chosen approach. Examples of papers, which discuss various methodological aspects, are Dyrberg (2004), Rommer (2005b), Rommer (2005c), Altman, Marco and Varetto (1994), Back, Laitinen, Sere and Wezel (1996), Begley, Ming and Watts (1996) and Frydman, Altman and Kao (1985).

Concerning the evaluation of results (item 3) credit institutions have to show that their models have discriminatory power, i.e. that the models can discriminate between defaulting and non-defaulting borrowers. In the credit-scoring literature it is common to report the type I errors (missing prediction, i.e. the model predicts a non-event, but it turns out to be an event) and type II errors (wrong signal, i.e. the model predicts an event, but it is a non-event). A good overview of the literature on validation, including how to assess the discriminatory power of a credit-scoring model, can be found in BCBS (2005). In this paper, we have chosen not to focus on the validation of the credit-scoring models, which we have estimated. Instead we wanted to keep the story simple and only discuss different credit-scoring models and their implications for the calculated capital requirements.[13]

---

[13] It is not simple to compare the discrimination abilities of the models, as the single-country models and the multi-country models are estimated using different portfolios. For further details the reader is referred to Hamerle, Rauhmeier and Rösch (2003).

Box 6: One possible reason for the differences in the calculated capital requirements: Different predictors of financial distress in the countries

That the predictors of financial distress are different in the countries is in line with Rommer (2005a), who compares the determinants of financial distress in French, Italian and Spanish small and medium-sized enterprises (SMEs) and concludes, that the estimation of single-country credit-scoring models show that there are some similarities among the predictors of financial distress across countries, but also that there are important differences.

Rommer (2005a) compares the significance and sign of the determinants of financial distress in the estimated credit-scoring models for the three countries. The comparison shows that the core variables that behave similarly across countries are the earnings ratio and the solvency ratio. They are significant and have a negative sign in all countries. A number of variables have effects that differ between the countries in terms of whether or not they are significant or what sign they have. These are the loans to total assets ratio, size, age, legal form and one of the ownership variables (very concentrated ownership). The differences in the significance levels in the countries may be due to a number of reasons, c.f. Rommer (2005a:26), who e.g. has the following explanation for why the legal form dummy (which is equal to one, when the legal status of the company is a private limited liability company and equal to 0, when the legal status of the company is a public limited liability company) is only significant in the Italian case, where it has the hypothesized positive sign: "The level of share capital between public and limited liability companies differ between the countries. In Italy the difference in share capital between the two types of legal forms is 110,000 euro, in Spain it is 60,000 euro and in France it is 37,000 euro… As only firms with 10 employees and a balance sheet of at least 2 million euro are considered in the estimations …, it is not surprising that only an effect of the private limited liability variable for the Italian firms, for which the difference in share capital between the private and public limited liability companies is the largest, is significant." Some variables are insignificant in both set-ups. These are the number of subsidiaries a firm has registered, the number of shareholders a firm has registered and two of the ownership variables (medium concentration and not very concentrated).

The result, which is obtained from the estimations of the single-country models, is confirmed by the estimation of a multi-country credit-scoring model (without country dummies). The estimations in Rommer (2005a) show that the multi-country model delivers parameter estimates that differ markedly from all the single-country credit-scoring models.

7. Conclusion

The Basel Committee's Revised Framework for Capital Measurement and Capital Standards (Basel II) will enter into force in 2007. Basel II facilitates the use of banks internal models to estimating probability of default when calculating the minimum capital requirement in the internal ratings-based approaches (IRB). Valid estimates of the probability of default require a considerable amount of data and default observations. Basel II allows for banks to pool their data to overcome their data shortcomings and a number of international data pooling projects have emerged. Thus even major international banks seem to need more data in order to fulfil the model requirements of Basel II.

To our knowledge, so far no study has compared the banks' capital requirements calculated on the basis of probability of defaults estimated from single-country credit-scoring models and multi-country credit-scoring models and discussed the incentive structure this might create for banks pooling data.

To illustrate the consequences on the calculated capital requirements of pooling data, we constructed a loan portfolio of loans to small and medium-sized enterprises for a hypothetical bank operating in France, Italy and Spain. For this purpose we use data

extracted from the Amadeus database provided by Bureau van Dijk. Using this data, the probability of default was estimated on the basis of single-country credit-scoring models and on the basis of multi-country credit-scoring models with pooled data from the three countries (with and without country dummies). The estimated probabilities of defaults are then used for calculating the minimum capital requirements for the hypothetical loan portfolio in the case of using country credit-scoring models, in the case of using a multi-country credit-scoring model with country dummies and in the case of using a multi-country credit-scoring model without country dummies.

Though our default definition is the same for the three countries and we controlled for variables such as age, size, legal form and sector, we find quite large differences in terms of the resulting minimum capital requirements for the portfolio in each of the three countries, when the probability of default is estimated using a single-country credit-scoring model compared to using multi-country credit-scoring models with and without country dummies.

One reason why the calculated capital requirements are different in the countries depending on whether the single-country or a multi-country model (with or without country dummies) is estimated could be that the predictors of financial distress differ across countries.

The results suggest that there might be incentives for cherry-picking, i.e. that banks choose a certain method because it delivers a lower capital requirement without considering what level of capital is actually appropriate to cover the overall credit risk. The overall calculated capital requirements vary with up to 18 percent depending on the choice of method for the hypothetical bank. Calculated for the individual countries it varies up to 47 percent.

Our hypothetical bank would obtain the lowest capital requirement from estimating a multi-country model with country dummies. In the situation where a bank considers pooling data with banks from other countries, the bank would also have an incentive to choose the method, which delivers the lowest capital requirement. The credit-scoring model, which delivers the lowest capital requirement, differs between the countries. For instance a bank with exposures to Italian firms would choose the single-country model, whereas a bank with the exposures to Spanish firms would choose the multi-country model with country dummies etc.

The calculations in this paper are based purely on the quantitative and technical requirements of Basel II and the EU directive proposal. We do not take into consideration effects of applying human judgment and a conservative perspective on PD estimation. The example of our hypothetical bank serves to illustrate the purely technical consequences of pooling data and the incentives it might give for banks when calculating minimum capital requirements.

The results are of particular interest for banks operating in different countries, which plan to pool data from their exposures in the various countries in order to estimate PDs like our hypothetical bank, maybe due to lack of a sufficient single-country database. The results are equally interesting for banks planning to pool data with banks from other countries to estimate PDs to make up for an insufficient database.

The overall conclusion from the analysis is that banks and supervisory authorities must be aware that the pooling of data from several countries should be done with caution. As there are not many official guidelines from authorities on the issue, we believe that the illustration in this paper of the consequences of pooling data from

several countries serves as an important input to the debate on how to set up credit-scoring models in banks that have cross-border exposures as well as for banks who choose to pool their data with banks in other countries. We have shown that it is not enough for banks to apply similar definitions of default and similar accounting regimes in the countries. Banks and regulators should also have a careful look into the models.

Literature

Altman, E. I., 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, vol. 23, no. 4, pp. 589-609

Altman, E. I. and A. Saunders, 1998. Credit risk measurement: Developments over the last 20 years. *Journal of Banking and Finance*, vol. 21, pp. 1721-1742

Altman, E. I., G. Marco and F. Varetto, 1994. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance*, vol. 18, pp. 505-529

Andersen, B. N., 2004. *Annual Meeting of the Danish Bankers' Association on 1 December 2004.* Speech by Governor Bodil Nyboe Andersen, Danmarks Nationalbank

Back, B., Laitinen, T., Sere, K. and M. v. Wezel, 1996. *Choosing Bankruptcy Predictors Using Discriminant Analysis, Logit Analysis, and Genetic Algorithms.* Turke Centre for Computer Science, Technical Report No. 40

Balcaen, S. and H. Ooghe, 2004. *Alternative methodologies in studies on business failure: do they produce better results than the classical statistical methods?* Working Paper 2004/249, Faculteit Economie en Bedrijfskunde

BCBS, Basel Committee on Banking Supervision, 1988. *International Convergence of Capital Measurement and Capital Standards.* Bank for International Settlements, July 1988

BCBS, Basel Committee on Banking Supervision, 2003. *Press releas: Significant Progress on Major Issues* . Bank for International Settlements, 11 October 2004

BCBS, Basel Committee on Banking Supervision, 2004. *International Convergence of Capital Measurement and Capital Standards. A Revised Framework.* Bank for International Settlements, June 2004

BCBS, Basel Committee on Banking Supervision, 2005. *Studies on the Validation of Internal Rating Systems.* Working paper no. 14, Basel Committee on Banking Supervision, February 2005

Beaver, W., 1966. Financial Ratios as Predictors of Bankruptcy. *Journal of Accounting Research*, vol. 6, pp. 71-102

Begley, J., Ming, J. and S. Watts, 1996. Bankruptcy Classification Errors in the 1980s: An Empirical Analysis of Altman's and Ohlson's Models. *Review of Accounting Studies*, vol. 1, pp. 267-284

Berger, A.N., Herring, R.J. and Szegö, G.P., 1995. The role of capital in financial institutions. *Journal of Banking and Finance*, vol. 19, pp. 393-430

BIS, 2004. *BIS holds Annual General Meeting and releases its 74th Annual Report*. Press release, June 2004

Caruana, J., 2004a. *Risk management trends and the supervisory structure*. Speech at the 31st General Assembly of the Geneva Association, Madrid, May 2004

Caruana, J. 2004b. *Basel II – emerging market perspectives*. Speech at Banker's Conference 2004, New Delhi, November 2004

Cass, D. (editor), 2002. Data hurdles. *Risk magazine*, vol. 15, no. 11

Dimitras, A. I., Zanakis, S. H. and C. Zopounidis, 1996. A survey of business failures with an emphasis on prediction methods and industrial applications. *European Journal of Operational Research*, vol. 90, pp. 487-513

Dunbar, N., 2005. Banks leapfrog the rating agencies. *Risk magazine*, vol. 18, no. 2

Dyrberg, A., 2004. *Firms in Financial Distress: An Exploratory Analysis.* WP no. 17, Danmarks Nationalbank

European Commission, 2004. *Proposals for directives of the European Parliament and of the Council, Re-casting Directive 2000/12/EC of the European Parliament and of the Council of 20 March 2000 relating to the taking up and pursuit of the business of credit institutions.* COM (2004), 486 final.

Fabi, F., Laviola, S. and P. M. Reedtz, 2004. *The treatment of SME loans in the new Basel Capital Accord: Some Evaluations.* Unpublished manuscript

Frydman, H., Altman, E. I. and D. Kao, 1985. Introducing Recursive Partitioning for Financial Classification: The Case of Financial Distress. *The Journal of Finance*, vol. XL, no. 1, pp. 269-291

UK FSA, Financial Services Authority, 2005. *Strengthening capital standards*

Greene, W., 2003. *Econometric Analysis*. New Jersey, USA: Prentice Hall

Greuning, H. van and Bratanovic, S.B., 2003. *Analyzing and Managing Banking Risk. A Framework for Assessing Corporate Governance and Financial Risk.* The World Bank, Second edition, 2003.

Hamerle, A., Rauhmeier, R. and D. Rösch, 2003. *Uses and Misuses of Measures for Credit Rating Accuracy.* Unpublished manuscript

Hayden, E., 2003. *Are Credit Scoring Models Sensitive With Respect to Default Definitions? Evidence from the Austrian Market.* Unpublished manuscript

Jenkins, S., 2003. *Survival Analysis.* Unpublished manuscript

Jones, F. L., 1987. Current Techniques in Bankruptcy Prediction. *Journal of Accounting Literature*, vol. 6, pp. 131-164

Keefe, D., 2004. *Caruana expects slower Basel II adoption*. Global Risk Regulator, November 2004

Kjeldsen, K., 2004. The Role of Capital in Banks. *Monetary Review*, Danmarks Nationalbank, 3rd Quarter.

Lando, D., 2004. *Credit Risk Modeling. Theory and Applications.* Princeton Series in Finance

La Porta, R., F. Lopez-De-Silanes, A. Shleifer and R. Vishny, 1998. Law and Finance. *The Journal of Political Economy*, vol. 106, no. 6, pp. 1113-1155

Masschlein, N., 2003. The Basel II Capital Accord, SME Loans and Implications for Belgium. *Financial Stability Review*, Nationalbank of Belgium, June

Mercer Oliver Wyman, 2005. Presentation at the *Seminar on Capital Allocation in Banks*. Joint training programme of Deutsche Bundesbank and Oesterreichische Nationalbank, Vienna, 27 to 28 April 2005.

Moody's Investors Service, 2001. *Moody's RiskCalc For Private Companies: Spain. Rating Methodology*, July 2001

Oestereichische Nationalbank, 2004. *Guidelines on Credit Risk Management, Rating Models and Validation*

Ohlson, J. A., 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, vol. 19, pp. 109-131

Rommer, A. D., 2005a. *A Comparative Analysis of the Determinants of Financial Distress in French, Italian and Spanish firms.* WP no. 26, Danmarks Nationalbank

Rommer, A. D., 2005b. *Testing the Assumptions of Credit-scoring Models.* WP no. 28, Danmarks Nationalbank

Rommer, A., D., 2005c. *Firms in Financial Distress: An Exploratory Analysis.* Unpublished manuscript

Saurina, J. and C. Trucharte, 2003. *The impact of Basel II on lending to small-and medium-sized firms. A regulatory policy assessment based on the Spanish Credit Register.* Bank of Spain, June

Shumway, T., 2001. Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, vol. 74, no. 1, pp. 101-124

Thoraval, P. and Duchateau, A., 2003. *Financial Stability and the New Basel Accord*. Banque de France, Financial Stability Review, no. 3, November 2003.

Appendix 1: Calculation of the minimum capital requirement for SMEs [14]

This appendix presents the formulas for calculation of the Basel II minimum capital requirement for SMEs.[15]

The formula specified by the Basel Committee for calculating the minimum capital requirement for a credit exposure (K*) is 8 % of the risk weight (RW) multiplied by the exposure at default (EAD):

$$K^* = 0.08 \times RW \times EAD$$

For exposures not in default, the formula under the IRB-approaches for calculating the risk weight (RW) is:

$$RW = \left[LGD \times N\left(\frac{N^{-1}(PD) + \sqrt{R} \times N^{-1}(0.999)}{\sqrt{1-R}} \right) - (LGD \times PD) \right] \times \frac{(1 + (M - 2.5) \times b)}{1 - 1.5 \times b} \times 12.5 \times \lambda$$

PD is the probability of default, LGD is the loss given default, and R is the assumed asset value correlation to systematic risk. N(x) denotes the cumulative distribution function for a standard normal random variable. The confidence level $N^{-1}(0.999)$ is set to 99.9 per cent. The first part in the squared brackets of the RW-formula is the assumed risk distribution of the losses, which is expressed as a function of LGD, PD and R.

The second part in the squared brackets ensures that the expected losses (-LGDxPD) are removed from the RW, as the minimum capital requirement under the IRB-approaches shall only cover unexpected losses. The constant 12.5 is the inverse of the capital requirement of 8 per cent. λ is a scale factor which was introduced by the Basel Committee to reiterate the Basel Committees objective of maintaining the current level of minimum capital requirements[16]. The Basel Committee has made it clear, that Basel II aims at the same global capital level as the 1988 Accord. The current best estimate of the scale factor from the Basel Committee is 1.06, c.f. BCBS (2004:14). The final determination of the scaling factor will probably be taken after the 5th quantitative impact study in 2005 and before the implementation of the Basel II, i.e. year-end 2006.

---

[14] The appearance of the formulas can seem a bit arbitrary, but one must bear in mind that the formulas are a result of economic and mathematical theory, several impact studies and not the least a pragmatic compromise between very different views and interests. The Basel Committee has chosen to be very brief in their explanation of the formulas. We will therefore not go into great detail explaining the formulas, also because the main focus in this paper is to illustrate the consequences of pooling data for corporate default risk by means of the Basel II capital requirement.

[15] The notation in the EU directive proposal (see European Commission (2004)) is slightly different from the notation in the Basel Committee's Revised Framework for Capital Measurement and Capital Standards (BCBS (2004)). We have chosen to follow the notation in the EU directive proposal.

[16] For further details of the reasoning for the introduction of the scale factor see the press release, c.f. BCBS (2003: 11 October 2003).

The last part of the formula is dealing with maturity effects. If the effective maturity (M) measured in years is equal to 2.5, the term in the squared brackets is reduced to a function of b:

$$b = \left(0.11852 - 0.05478 \times \ln(PD)\right)^2$$

The purpose of b is to transform the one-year time horizon, which is the time horizon for PD, to a "longer maturity" minimum capital requirement.

The correlation to systematic risk R is determined by:

$$R = 0.12 \times \frac{1 - e^{-50 \times PD}}{1 - e^{-50}} + 0.24 \times \frac{1 - \left(1 - e^{-50 \times PD}\right)}{1 - e^{-50}} - \omega^{SME}$$

The R-function is an estimate of the link (correlation) between the joint default of two separate borrowers. The IRB model relies on a single-factor asset value model to describe the co-movement of defaults in a portfolio. The single-factor can be interpreted as a variable, which represents the state of the economic cycle. IRB correlations to the single-factor are a decreasing function of the borrower's credit quality PD. The best credit quality borrowers (with a small PD) have a correlation of 24 %, and the lowest credit quality borrowers (with a high PD) have a correlation of 12 %, c.f. the R-formula.

For exposures to SME borrowers R is also a function of the firm size ω^{SME}:

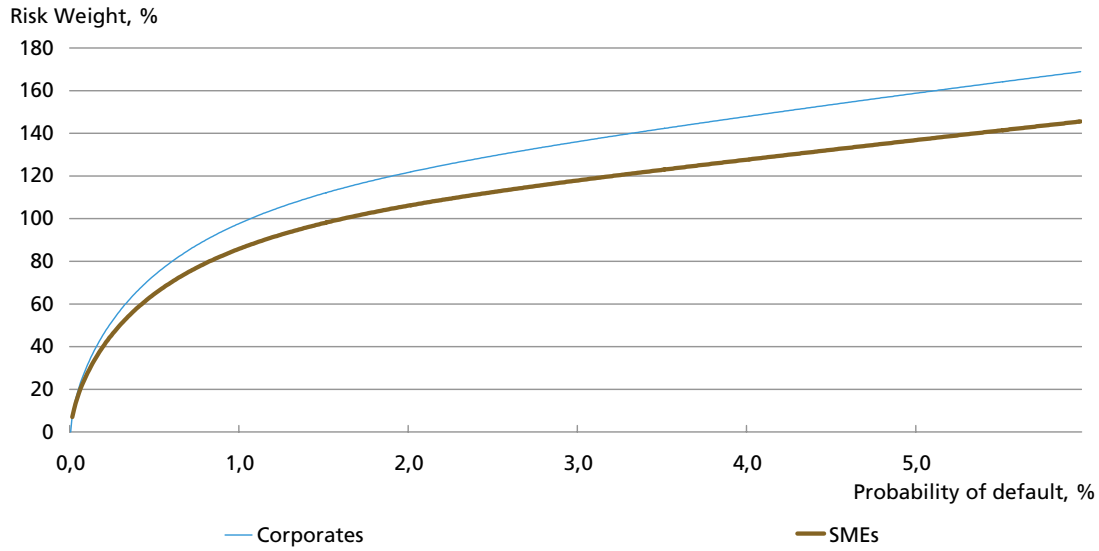
$$\omega^{SME} = 0.04 \times \left(1 - \frac{S - 5}{45}\right)$$

S is expressed as total annual sales in millions of euros for the companies, which have total annual sales between 5 and 50 million euro. Companies, which have reported sales under 5 million euro, will be treated as if they had sales of 5 million euro. The capital reduction increases linearly from 0 to 20 per cent with sales going from 50 to 5 million euros, and remains at 20 per cent for firms with sales figures lower than the latter threshold.

The Basel Committee has decided that the PD for corporate exposures (including SMEs) has to be larger than or equal to 0.03 %. This means that if the estimated PDs are less than 0.03 %, it should be set to 0.03 %, when calculating the risk weight. Under the foundation IRB approach, the Basel Committee has set the effective maturity (M) to 2.5 years and LGD for corporate claims to 45%

As illustrated in chart A1 the risk weight for exposures to SMEs are lower than for the exposures to corporates, and the difference increases with the size of PD.

Chart A1: IRB-curves for corporate and SME exposures



Note: LGD=45 %, M=2.5 and S= 25 million euro.

Appendix 2: Results

Table A.2.a: Results based on the PDs from the single-country credit-scoring models

	Number of active companies in 2002 in the data set	The average probability of default for the companies that were active in 2002	The average capital requirement for the companies that were active in 2002, in 1000 euro
IT	35818	0.00138	2644
ES	29447	0.00135	4163
FR	41251	0.00544	2425

Note: The probability of default for each company is calculated as the average probability of default for the period 2000 – 2002. For a specific company the average probability of default can be calculated as the average of 1, 2 or 3 data points, depending on when the firm entered the sample.

Table A.2.b: Results based on the PDs from the multi-country credit-scoring model (without country dummies)

	Number of active companies in 2002 in the data set	The average probability of default for the companies that were active in 2002	The average capital requirement for the companies that were active in 2002, in 1000 euro
IT	35818	0.00282	3899
ES	29447	0.00257	5051
FR	41251	0.00328	1985

Note: The probability of default for each company is calculated as the average probability of default for the period 2000 – 2002. For a specific company the average probability of default can be calculated as the average of 1, 2 or 3 data points, depending on when the firm entered the sample.

Table A.2.c: Results based on the PDs from the multi-country credit-scoring model (with country dummies)

	Number of active companies in 2002 in the data set	The average probability of default for the companies that were active in 2002	The average capital requirement for the companies that were active in 2002, in 1000 euro
IT	35818	0.00154	2773
ES	29447	0.00144	3730
FR	41251	0.00520	2533

Note: The probability of default for each company is calculated as the average probability of default for the period 2000 – 2002. For a specific company the average probability of default can be calculated as the average of 1, 2 or 3 data points, depending on when the firm entered the sample.

Table A.2.d: The exposure of default (EAD) in the countries

	Number of active companies in 2002 in the data set	The average exposure at default in the companies that were active in 2002, in 1000 euro
IT	35818	1140
ES	29447	1611
FR	41251	556