ANO 2004/2

Oslo
February 26, 2004

Working Paper

Research Department

Aggregate bankruptcy probabilities and their role in explaining banks' loan losses

by

Olga Andreeva

NORGES BANK

Working papers fra Norges Bank kan bestilles over e-post:
posten@norges-bank.no
eller ved henvendelse til:
Norges Bank, Abonnementsservice
Postboks 1179 Sentrum
0107 Oslo
Telefon 22 31 63 83, Telefaks 22 41 31 05

Fra 1999 og senere er publikasjonene tilgjengelige som pdf-filer
på www.norges-bank.no, under “Publikasjoner”.

Working papers inneholder forskningsarbeider
og utredninger som vanligvis
ikke har fått sin endelige form.
Hensikten er blant annet at forfatteren
kan motta kommentarer fra kolleger
og andre interesserte.

Synspunkter og konklusjoner i arbeidene
står for forfatternes regning.

Working papers from Norges Bank can be ordered by e-mail:
posten@norges-bank.no
or from Norges Bank, Subscription service,
P.O.Box. 1179 Sentrum
N-0107Oslo, Norway.
Tel. +47 22 31 63 83, Fax. +47 22 41 31 05

Working papers from 1999 onwards are available as pdf-files on the bank’s
web site: www.norges-bank.no, under “Publications”.

Norges Bank’s working papers present
research projects and reports
(not usually in their final form)
and are intended inter alia to enable
the author to benefit from the comments
of colleagues and other interested parties.

Views and conclusions expressed in working papers are
the responsibility of the authors alone.

ISSN 0801-2504 (printed), 1502-8143 (online)
ISBN 82-7553-225-6 (printed), 82-7553-226-4 (online)

Aggregate bankruptcy probabilities
and their role in explaining banks' loan losses

Olga Andreeva
26 February 2004

Abstract

Increased competition forces banks to narrow lending margins and at the same time relaxed lending standards worsen the pool of borrowers. To preserve sound banking system it is important task to monitor credit risk as one of the dominant factors leading to bank failures and financial vulnerability. Norwegian banks traditionally have a large share of loans to non-financial enterprises in their investment portfolios, and we focus on risk related to loans provided to limited liability enterprises. By combining statistics on loans to Norwegian industries and regions and bankruptcy probabilities for individual corporate borrowers, we construct a proxy reflecting risk profile of the banks' loan portfolios. Aggregation within industries and counties provides a bank-level panel of risk indicators, which are used to estimate banks' loan losses during the period 1988 – 2001. Constructed aggregate bankruptcy probabilities prove to be meaningful measures, which explain loan losses if we control for the macroeconomic and bank specific factors.

JEL Code: G21, C81
Key words: Bank losses, bankruptcy probabilities, aggregation

Acknowledgement: I would like to thank Bent Vale, Kjersti-Gro Lindquist, Glenn Hoggarth and participants of the seminars in the Research Department and Financial Stability Wing for valuable comments and discussions.

1. Introduction

One of the most important roles of banks as financial intermediaries is allocation of credit, screening and monitoring of borrowers' creditworthiness, and maintaining relationships with reliable customers, which they can do on a lower costs than individual agents. Bank loans are especially valuable for small firms that are not publicly traded and thus are constrained with financial resources due to the limited access to the financial markets.[1]

Well functioning financial markets and market discipline play an important role for preserving soundness of the banking system and keeping risks in adequate limits. However, market failures, free-rider problems of gaining benefits from collected information and other forms of distorted incentives of economic agents advocate for the presence of sound regulation.[2] The New Basel Capital Accord also emphasises supervisory review process as an important part of controlling risks in banking.

Credit risk and financial stability

Financial system is exposed to four major types of risks related to the financial intermediaries: liquidity risk, market risk, credit risk, and operational risk. One of the central issues of the financial stability reports is to measure and monitor these risks, examine risks patterns and assess financial system vulnerability to them. Risk control policy is especially important in banks, the largest part of financial intermediaries, as bank failures induce large costs on the economy, society and government.[3]

It is widely recognised that credit risk is one of the dominant factors leading to bank failures and financial vulnerability. Lending is a main function of universal commercial banks and is even more inherent to savings banks, which allocate almost all attracted deposits to loans. Moreover, other types of risk reinforce credit risk to some extent, as for instance, due to the interest rate movements and changes in operational environment with counterparties bank may be exposed to higher credit risk.

Banks may take excessive risks due to various factors from intentional risk taking and high risk tolerance in a competitive environment in situations of moral hazard and adverse selection.[4] Even banks that apply good risk measurement techniques can underestimate potential risks due to low-frequency and high-severity event which may produce huge but almost unanticipated losses. As it is emphasised in Herring (1999), banks are often influenced by a special form of financial vulnerability, disaster myopia, when they undervalue default probabilities if failures do not arise for a long time. And even if a bank uses superior credit risk models that indicate higher risk pricing, it may lose in competition to other banks, which disregard this risk and therefore may choose herding behaviour. Increased competition from credit markets forces banks to narrow spreads and at the same time relaxed lending standards worsen the pool of borrowers.[5] Strong competition with disaster myopia, short termism and herding may therefore increase financial vulnerability of banks.

---

[1] Diamond (1991a), Becketti and Morris (1992).
[2] Financial Stability Review, Bank of England (2000-2002).
[3] See in more details in Mailath and Mester (1994), Frydl (1999), Hoggarth, Reis and Saporta (2002).
[4] See Mishkin (1991) on a discussion o asymmetric information and agency costs as causes of financial instability.
[5] See a discussion in Salas and Saurina (2002a) and Matutes and Vives (2000) on risk taking behaviour of banks as a response to changes in competition and market power.

Market discipline is also diminished by insured liabilities of the banks since banks depositors are secured and thus have less incentive for control. Bank assets can be easily misallocated as they can borrow easier and therefore take higher risks in asset allocation. Since sound banking and financial health are essential factors for financial stability, it is important to monitor bank risk exposure to the corporate sector, changes in lending patterns and ensuing losses.

Credit risk and loan losses

Most of the borrowers on the credit market have limited liability on their obligations to the bank and therefore lenders are exposed to the risk of borrowers default. Problem loans are one of the major reasons of financial difficulties, especially for banks with a large scale of traditional lending activities. To insure themselves, at least partially, from the borrowers' failure to repay, banks set aside loan loss provisions for expected losses on doubtful debts. Bank practices differ with respect to the rules used in definition of expected losses and estimations of loan loss provisions. Norwegian practice defines expected losses as losses inherent in the loan portfolio but not yet realised, and therefore loss provisions are based only on the current information. However, expected losses may also be defined as all possible future losses that can occur due to both current and future events, and thus indicate how much loss provisions a bank can make to account for possible future losses. Making such loan loss provisions, banks can write off losses against them and thus reduce the risk of weaker profitability and capital adequacy when losses are recognised. Systematic under-provisioning policy exposes bank credit portfolio to additional risk, as the bank may be unprepared to withstand shocks and maintain solvency.

At the same time, variation of losses is uncertain, and therefore unexpected loss should also be considered a possible danger for bank financial situation that increases the probability of insolvency, especially if the bank does not maintain sufficient capital in relation to its assets. Uncertain magnitude of possible losses gives rise to the credit risk. While loss provisions may cover expected losses on loans, bank capital in excess of the required minimum helps to absorb unexpected losses so that a bank can maintain solvency.

When banks decide on their lending policy they have a trade-off between short-term gain from risk-taking and long-term losses on loans and possible bankruptcy or takeover. Considered costs and losses also include expected loss, assessment of its possible variability and opportunity cost of allocating capital and liabilities. Expected loss can be calculated on the basis of borrowers' creditworthiness and correlation of loss exposure of different loans in the portfolio. If allocation of credit is not profitable, a bank may increase interest rate on loans or collateral requirements to reduce expected loss if it cannot reduce costs. However, this policy is not always sustained due to the downward competition press on interest rates.

Approaches to credit risk and motivation for the study

Due to the common concern of regulators in many countries about the financial stability a lot of effort has been done in the direction of assessment of credit risk and construction of warning indicators based on these measures. Credit risk is associated with the possibility that the borrower will not fulfil its contractual obligations and depends on the general macroeconomic situation, lending standards, i.e. interest rate, collateral requirements and other loan covenants, and legal enforcement mechanism, including the capacity to recover part of the loan after the default. There exist many different approaches to measuring credit risk and assessing its influence on bank performance. Value at risk models (VaR), option-based and insurance approach[6] to risk measurement and also rating-based models try to quantify credit risks and exposures of the banks. The size of risk is measured as the amount of a potential loss that can be incurred by a bank with some probability. Some of the models are designed on quite a sophisticated level and they often require extensive data for different contingencies and even confidential information related to the banks' internal accounts and customers' financial position. Lack of this information or low quality information can widely decrease supervisory effects from these models.

A natural approach to the credit risk measurement when credit claims are not tradable is to measure a probability of default to occur and amount of loss given that default. Loss in the event of default is the amount of money that the bank will not be able to recover less possible recoveries on collateral. Then expected loss is a probability of default over the next year multiplied by the loss given default. But accurate estimation of the default probabilities requires quite detailed information on borrowers.

Norwegian banks are mainly engaged in traditional banking with loans constituting the largest part of their assets. Therefore, we concentrate on a narrow meaning of the credit risk, i.e. risk related to bank loans. The aim of the analysis is to construct a proxy for the credit risk measure to reflect risk profile of the banks' loan portfolios. In order to do this we aggregate risk indicators for banks on the basis of bankruptcy probabilities for individual corporate borrowers[7], and estimate how these indicators can explain banks' loan losses during the period 1988 - 2001. Two types of annual data are combined for this study: detailed bank statistics on loans specified for each county and industry and statistics for individual non-financial enterprises with limited liability. To construct a risk measure for a bank, bankruptcy probabilities for enterprises are aggregated within county and/or industry groups and then weighted by the volume of loans granted to each of these groups by this bank. Commercial banks have higher share of corporate loans, while savings banks traditionally provide loans mostly to households. However, historically mortgages are safer than loans to corporations (within the present and the New Basel Capital Accord house mortgages are also considered less risky), therefore we do not lose much by focusing on industrial loans in our risk assessment. Constructing a risk measure for the banks' loan portfolios which can explain bank loan losses is an important task in studying the banking system and preserving its soundness.

2. Description of the datasets

Statistics on bank loans

We consider annual aggregate volumes of domestic loans of the Norwegian savings and commercial banks and branches and subsidiaries of foreign banks in Norway to the non-financial institutions classified by industry and county.[8] The number of Norwegian banks is gradually decreasing from around 150 savings banks and 20 commercial banks at the beginning of the sample period to 130 and 12 banks respectively in 1999/2000. At the same time, volume of loans adjusted for the Consumer price index (CPI) index is generally growing with exception of 1990-1991 and 1993-1994. The data in its most disaggregated form is represented by loans

[6] See Saunders (1999) on VaR, KMV, insurance and other approaches to credit risk measurement.
[7] See Bernhardsen (2001) and Eklund, et al (2001) for estimation of individual bankruptcy probabilities.
[8] Information is taken from the banks financial reports (Report 60). Data on loans granted by other financial enterprises and mortgage companies, which constitute almost 40 per cent of all observations (around 20 per cent in volume of loans), are available only from 1996 and are not included in the data set.

to around thirty – sixty industries[9] and nineteen counties[10] because information on the individual borrowers of each bank is not available. According to this type of classification we combine data from the banks' end of year balance sheets with annual statistics on individual enterprises along two dimensions: industry dimension and industry/county dimension. Later they are refered to as industry/year and industry/county/year groups.[11] We use only the data on loans granted by banks to the sector of limited liability enterprises over the years 1988 - 2001.

Data was controlled against negative observations for loans and positive observations for loan loss provisions. Observations with missing or zero industry and county codes were dropped.

Statistics on enterprises (annual financial statements): SEBRA-database

The SEBRA-database is a broad dataset on limited liabilities enterprises. We have excluded companies in the oil and gas industry, financial industry and public sector. It contains information from annual financial statements of the enterprises registered at the Norwegian register for business enterprises over the years 1988-2001. The data set contains 1,399,119 observations in total for 14 years. The number of enterprises submitting their financial records was constantly growing from 47,641 in 1988 to 137,201 in 2000 with a small decrease in 1994, but there is a large drop of more than 7 per cent in the last period of the data set, year 2001. At the same time, number of enterprises in different industries and counties varies from just a few to several thousands. This pattern is similar to the statistics on loans, which can be explained by a relatively low level of activities in some counties and industries. The dataset was checked for missing observations for those enterprises that provide accounting information not on a regular basis. The data was controlled against missing and zero industry and county codes, and also against observations with industry codes that do not correspond with aggregate codes in the bank statistics.

The SEBRA model[12] predicts bankruptcy probabilities for individual enterprises with book value of total assets exceeding 250,000-300,000 NOK on the basis of accounting statements. An observation is defined as a record with financial and other relevant information submitted by an enterprise (referring to its unique identification number) available in the database for a particular year. High average bankruptcy probabilities with large deviations, i.e. mean value larger than 0.036 and standard deviation larger than 0.065, which corresponds to the upper 25 per cent, are found in many industries especially during the Norwegian banking crisis years 1990-1993. High bankruptcy probabilities during the years beyond the crisis are found in the following industries: Fishing, Manufacture of office machinery and computers, Hotels and restaurants, Post and telecommunication, Recreation, cultural and sporting activities, Other service activities. These industries traditionally have high uncertainty in their activities, which is particularly true for the hotel, restaurants, recreation, service activities and fishing. However, Real estate activities, which are also considered risky, show quite stable and low values of bankruptcy probabilities throughout the sample period.

[9] Standard classification includes 32 industries before 1991, 33 industries up to 1996, 58 industries in 1996-1997 and 59 industries up to 2001.

[10] Observations for counties 21 – 23 were joined in county 21 (Svalbard) as counties 22 and 23 are not defined in the enterprise statistics, and observations for county 2 (Akershus) and county 3 (Oslo) were joined in county 3 (Oslo/Akershus) due to the geographical and economic interrelations of these counties.

[11] Since we use data classified by industry, changes in the type of industry classification in the bank reports (i.e. the number and contents of specified industries) can explain the variation in the number of groups (e.g. introduction of a more detailed classification in 1996 gives a rise in the number of observations to more than 6,400 compared to around 4,200 in the previous years).

[12] See Bernhardsen (2001) and Eklund, et al (2001) for a description of the model.

Linking of the datasets and aggregation of individual bankruptcy probabilites
The SEBRA-database contains only industry codes consistent with SIC94 as they were previously converted from SIC83 for all enterprises, while the bank statistics use old aggregate classification of industries in Reports 60 up to 1996. Therefore, for the data before 1996 we assign old aggregate codes to enterprises using relationship patterns between old aggregate codes and SIC83, and between SIC83 and SIC94. For the data from 1996 to 2001, assignment of the aggregate industry codes, valid in the bank statistics after 1996, to enterprises in the SEBRA-database is made according to the relationship pattern between SIC94 and aggregate codes in the Report 60. In this respect, a formal correspondence pattern between two industry classifications is utilised, where possible; whereas some artificial relationship between them is suggested, where necessary.[13]

After establishing a correspondence between industry codes in the bank statistics and industry codes for the individual enterprises, we aggregate individual bankruptcy probabilities, obtained for each enterprise from the SEBRA-model. Referring to the two common dimensions for the banks' reports and the SEBRA-database, we use industry/year groups, i.e. the aggregate across all counties, and industry/county/year groups. The first type of aggregation mixes observations across counties and can be in disagreement with the county specific type of activities of the medium-size savings banks. However, it provides a direct link between the two datasets. Moreover, it may be more accurate than the second one if banks in their annual reports assign counties on some other basis (e.g. location of the local branch which an enterprises uses for its loan application), than the formal registration criteria used in the SEBRA-database. The second type of aggregation allows utilisation of higher variation in risk indicators, i.e. over larger number of groups. Volumes of debt to the financial institutions or the levels of activities, represented, for example, by total assets or operating revenues are used as weights in aggregation. It is reasonable to focus on the enterprises with non-zero 'debt in financial institutions' since only these enterprises will inflict a loss for the bank in the event of bankruptcy.

Probability of non-repayment of the loan may depend on the borrowers' prospects and type of business as well as financial strength and liquidity characteristics. These factors are incorporated into the bankruptcy probabilities through financial ratios reflecting companies' earnings, liquidity and solidity, as well as companies and industry characteristics (age, size, and deviations of the profitability, liquidity and solidity from industries averages).[14] Therefore, aggregated bankruptcy probabilites serve as a good risk indicator and can be used to estimate loan losses for individual banks.

However, we do not have a direct link to the borrowers of each bank and also financial information is subject to a quick change, which creates a scope for upward or downward biases in loan losses estimation based only on these risk measures. So banks' risk profile is not completely reproduced and when we model bank loan losses we need to incorporate some proxies for distinguishing between banks' lending policies. Therefore, we consider also macroeconomic data, interest rate, and some bank-specific information which is discussed below.

[13] See a detailed description of these procedures in the appendix "Combining bank statistics on loans with statistics on non-financial enterprises".
[14] See Bernhardsen (2001) and Eklund, et al (2001).

3. A problem description

Loan losses vs. loan loss provisions

Loan losses consist of actual losses and changes in loan loss provisions, which are carried to reflect more accurate current value of bank assets. Specific loan loss provisions (tax deductible) are made on the specific loans which are identified as doubtful. General loan loss provisions (not tax deductible) are made solely to cover losses which can occur on the basis of the economic perspectives and industry analysis, when specific doubtful loans are not possible to identify.[15] A bank, which has had an adequate provisioning policy, writes off recognised losses on a loan against the stock of previously made loss provisions on this loan. If loan loss provisions are not made, actual losses are contributing directly to the increase in recorded (book) loan losses and may decrease current profitability (see Table 1 below). Therefore, loan loss is a measure of ex post credit risk.

Table 1: Loan losses and provisioning practice in the Norwegian banks

Actual losses not covered by previous loss provisions (write-offs)
+ Specific loan loss provisions on new loans + Net increase in specific loan loss provisions on previously made loans (increased provisions minus write-backs) + Increase in general loan loss provisions - Recoveries of previously written off loans losses + Other corrections
Recorded loan losses

Loan loss provisioning practice may vary across the banks due to different assessment of the borrowers financial conditions and performance, bank risk profile and corresponding practice of loan loss provisioning as a share of problem loans, collateral valuation and its role in reducing actual loan losses, and timing of writing off actual loan losses.[16] Moreover, as the size of timing and amount of the future actual loss is unknown provisions are subject to expectations which can be better during economic upturns and worse during downturns. So improving economic situation may lead to the reversals in provisions, while during a crisis banks may increase their provisions to a large extent.

Loan loss provisions my have a signalling effect. For example, Thakor (1987) discusses effect of assets write-downs in signalling forthcoming events and Musumeci and Sinkey (1990) claim that by making loss provisions banks not only adjust their accounting records according to the past events but also provide additional positive information to the market. Therefore, banks may conduct provisioning policy taking into account not only the amount of doubtful loans but also signalling effects. However, Scholes, Wilson and Wolfson (1990) find that if the market already had a good estimate of the bank's assets and earnings, then we could not expect any further effect on them by provisioning decisions. Moreover, as accounting rules for loan loss provisions are quite strict it is easier to write them back than to delay, while write-offs are more

[15] See Chirinko and Guill (1991) for the estimation of the portfolio risk dependent on the exchange rates, commodity prices, taxes, spending policies and regulation. Assessment of the exogenous portfolio risk is made on the basis of industries' performance, using proportion of each industry in portfolio and loan loss distribution for each industry.
[16] See Beattie et al. (1995) for a detailed discussion of current practices and alternative approaches to loan loss provisioning in banks.

discretionary as they are made when the loan is irrecoverable and is not expected to be repaid. Therefore, one would expect provisions to have less negative signalling effect than write-offs.

At the same time, specific loan loss provisions are made against equity capital and thus addition to them increases the cost of bank capital. Unanticipated large increase in loss provisions may therefore negatively influence bank's cost of funds and share price. This hypothesis is opposite to the one corresponding to the positive market reaction to the loss provisioning. However, the stronger is a bank's capital position the more easily it can undertake large loss provisions. Liu and Ryan (1995) show that loan loss provisions convey both positive and negative information to the marker depending on the loan portfolio composition. They found that market reaction to the increase in loss provisions for large and frequently renegotiated loans (i.e. commercial loans) is positive and for the increase in loss provisions for small and infrequently renegotiated loans (i.e. consumer loans) it is negative.

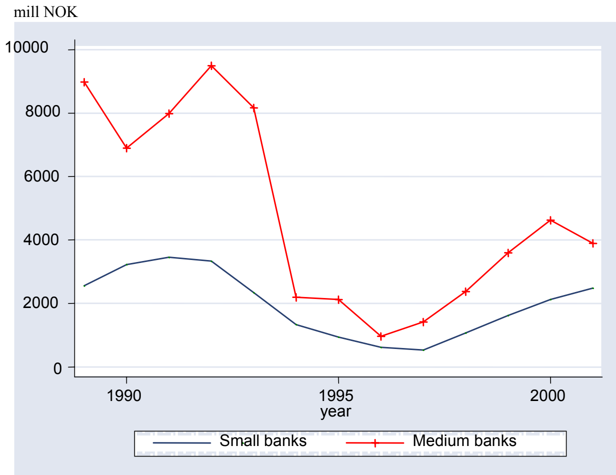
In general banks have an incentive to avoid showing losses that would imply reduction in capital as it may convey a negative signal to the market. Instead they can set interest margins to cover expected risks. However, intense competition may prohibit them from setting high interest margins on loans, and inexperienced lenders may intentionally or even unintentionally underprice.

Data features

Bank loan losses[17], stock of loss provisions and non-performing loans exhibit different patterns for small and medium versus large banks. The data show very low after crisis loan losses especially at the large banks which made large reversals of previously recorded losses and loan loss provisions. A rise in loan losses during the last years is also quite noticeable in contrast to previous reversals. Then these banks have started to make provisions on new loans and also to write off losses that were not covered by previous loan loss provisions.

Bank loan losses (sample period 1988 – 2001)

Small and medium size banks



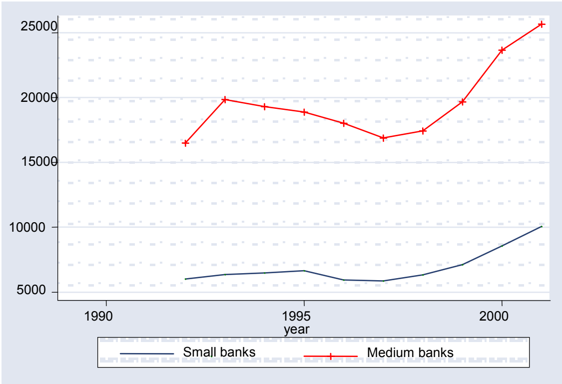
Large banks

[17] In the data and econometric analysis we consider recorded loan losses as defined in the Table 1 above.

Loan loss provisions

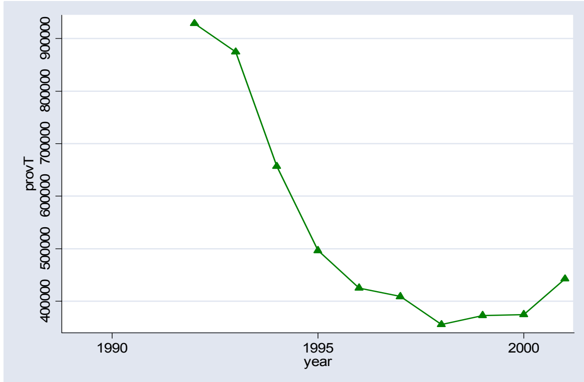
Small and medium size banks

mill NOK



Large banks

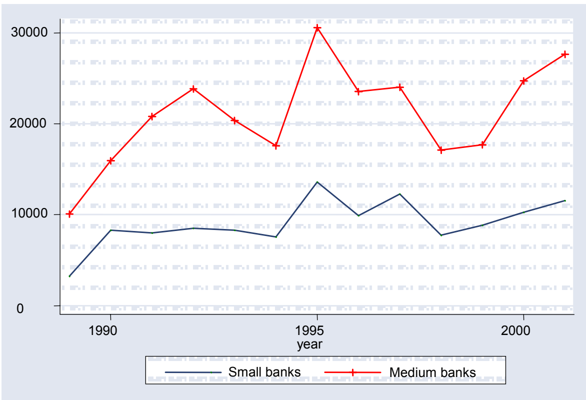
mill NOK



Non-performing loans

Small and medium size banks

mill NOK

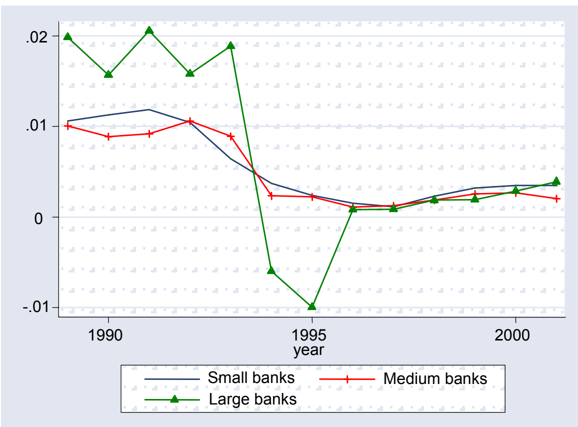


Large banks

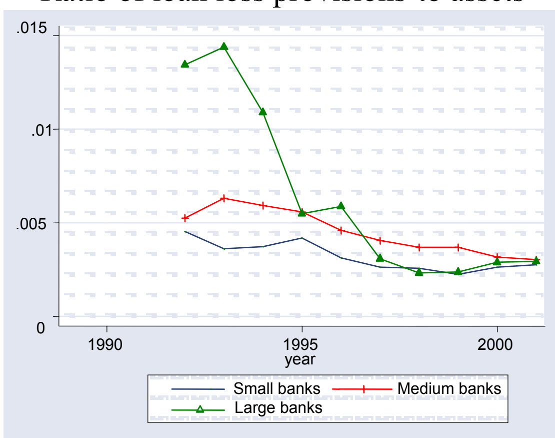
mill NOK

Recent reduced reversals of loss provisions and growth of loan portfolios lie at the basis of recent increase of loan losses, while measured in relation to gross loans, losses are not increasing dramatically. On the following graph we can see the patterns of the ratios of loan losses to assets and loan loss provisions to assets in all three groups of banks:

Ratio of loan losses to assets



Ratio of loan loss provisions to assets



The pattern of loan losses was very high for most of the banks during the crisis years and started to grow again in 1996, while the ratio of loan losses to assets exhibits a flatter pattern, partially due to the high growth in assets values.[18] The same is true if we compare patterns for loan loss provisions and the ratio of loan loss provisions to bank assets. Average loan losses constitute 35,596 mil NOK with total variation of around 276,533 mil NOK, where 150,000 mil NOK is standard deviation between the banks and 239,156 mil NOK is standard deviation over sample years. For the ratio of losses to bank assets with average of 0.0057 and standard deviation of 0.0167 we have closer values of between and within variation of around 0.0155 with a bit higher variation between banks. Similar pattern is seen for the ratio of loss provisions and non-performing loans to assets.

4. Motivation for the model

The aim of the current study is to build an econometric model allowing to test the quality of the constructed aggregate bankruptcy probabilities and to analyse how increase in the bank risk profile will enhance loan losses. Aggregate bankruptcy probability is our main testing variable, which is a proxy measure of risk for banks' loan portfolios. Financial strength of individual enterprises lies at the origin of aggregate bankruptcy probabilities, as enterprises with healthy financial accounts are able to absorb shocks and survive losses without going bankrupt. Worsening of the enterprises' financial situation leads to a higher probability of future bankruptcy.[19] The purpose of aggregating individual bankruptcy probabilities of enterprises with loans in financial institutions is to arrive at bank-specific information on risk associated with the portfolio of corporate loans.

Enterprises actual bankruptcy rate can be seen as a good economic indicator for predicting bank loan losses. However, it tends to develop with a lag of several years to the business cycle as actual bankruptcies are usually registered with a delay after the point when the firm cease to fulfil its financial obligations. Moreover, actual bankruptcy rate may not be helpful, as banks tend to make loss provisions on doubtful loans and write off irrecoverable loans, and therefore loans to a firm going bankrupt may already have been recorded as losses or written off. At the same time, the SEBRA model predicts the probabilities of bankruptcy happening in the following three years on the basis of the information available up to the current year. However, banks revise their credit policy in the current year based on the available public information, i.e. for the previous year, and therefore we take bankruptcy probabilities with a one period lag in the model. For example, if some enterprises experience worsening of their financial situation they may have problems with repayment of loans, which in turn leads to an increase in the size of non-performing and consequently loss provisions. Even non-performing loans themselves is a good indicator showing the tendency in loan losses, and hence it may be a measure that can add additional information in estimation of loan losses. However, banks have some discretion in their provisioning policy and some additional factors may influence the choice and assessment of doubtful and non-performing loans, and the extent of their provisioning. Therefore, extra information is needed in predicting the size of bank loan losses.

We build the analysis on a simple reduced form model and the following framework for estimation of bank loan losses is considered. The panel data regression analysis is used to test

[18] See Boyd, Gomis, Kwak and Smith (2001), and also Steigum (2004) for a discussion on the specific features of the Norwegian banking crisis.

[19] Alternative approach may be based on market evaluation reflecting expectations about enterprise future earnings, but it is only for publicly quoted firms.

the effectiveness of the risk measures constructed for each individual bank. We test how they can explain banks' loan losses controlling for macroeconomic and various bank specific factors. We incorporate in the econometric model some major economic factors to measure influence of each factor on the expected losses given other things constant.

The size of loan loss provisions/loan losses responds to the changes in risk proxied by factors directly related to the banks' loan portfolio and factors reflecting the general macroeconomic situation.[20] We have to check whether constructed risk measures for the banks' loan portfolio can explain variation in loan losses and how well they can contribute along with other factors as GDP, unemployment, housing prices and interest rates. At the same time, banks may experience different levels of caution in making loan loss provisions depending on their attitude to risk and overall ability to withstand unexpected losses and macroeconomic shocks. An important series of factors in explaining loan losses is therefore related to the bank-specific information. Bank specific indicators can be based on the two related factors: quality of banks management, i.e. quality and costs of the procedure of assessment, selection and monitoring of borrowers, preciseness in the estimation and pricing of expected risk; and quality of the current loan portfolio.[21] The second can decrease due to the deterioration of the borrowers' performance with time, including influence of macroeconomic shocks. Therefore, a measure of the probable default on the bank's portfolio of loans can be a useful indicator of the bank's credit risk, as it reflects the quality of the current borrowers and also indirectly incorporates some influence of worsening macro conditions.

What is particularly essential in our case is that aggregate bankruptcy probabilities reflect this information only partially. Credit risk is measured here with respect to the loans to different industries and regions but without a direct reference to a bank-specific client base. Aggregate bankruptcy probability reveals information only on the average quality of borrowers from a specific industry and region, and therefore reflects only the average risk for each bank due to its specialisation in particular industries and regions. But these risk measures do not take into account bank's individual customers and consequently variation in bankruptcy probabilities inside industry/region groups. Moreover, some banks may end up with worse borrowers than other either by chance or due to poorer risk management (i.e. fail to evaluate borrowers, to monitor their performance, to evaluate collateral properly) and higher risk aversion. For that reason we need to have some proxies reflecting banks' attitude to risk and quality of their management.

Additional effects

Residual variation in loss rate is very large and to reduce these shortcomings of bankruptcy probabilities that cannot explain much of the variation in bank loan losses, we incorporate macro and micro factors that influence bank loss rate.

<u>Macroeconomic trends</u> have a large impact on the pattern of loan losses. For example, compensation for risk in lending depends on the business cycle and bank's expectations about

[20] Fernandez , Martinez and Saurina (2000) study cyclical behaviour of bank loans, loan losses and loan loss provisions in Spain and show that housing prices, asset prices and lending margins have good explanatory power for bank lending.

[21] DeYoung (1997), Berger and DeYoung (1997) relate problem loans and bank efficiency considerations and argue that low quality banks with poor management may badly monitor not only borrowers but also costs.

future earnings prospects. GDP pattern is a good proxy for the position in the economic cycle and can serve as an additional explanatory variable for bank loan losses.

As it was discussed above, bank credit risk and consequently loan losses are mainly connected to the developments in the corporate sector. However, enterprises depend on the stable demand from the household side. Households are particularly vulnerable to the changes in their disposable income, which can be proxied by *unemployment rate*. In addition, changes in the interest rate in the economy, which influence interest rate on loans, also have some effect on the size of debt burden and thus vulnerability of households to economic changes, including unemployment rate. Both factors have a direct effect on the household debt-servicing capacity as it changes debt burden. The latter weakens households' ability to withstand macroeconomic downturns and worsens consumption capacity. Lower disposable income as a consequence of unemployment or growing interest rates for servicing the debt may therefore lead to a serious reduction in private consumption. The latter affects sales of most enterprises and decreases their debt-servicing capacity.

By this we have a two-sided effect of weaker household economy on the banks' loan portfolios. First, there is a direct effect through loans to households, as they may have higher difficulties in debt servicing while their debt burden increases. But also there is an indirect effect through corporate loans as the situation in the corporate sector is worsening due to the lower demand, which can lead to industrial loan losses without any significant rise in losses on household loans. Thus, unemployment variable reflects not only general macroeconomic changes but can also partially proxy credit risk associated with loans to enterprises and households. At the same time, higher share of household loans exposes banks more to household financial situation and risk of changes in the housing market through the collateral value.

Property prices

Housing prices reflect risks related to mortgages. Demand for houses, which boost the price, depends on households' disposable income, employment situation and interest rate on loans. An after-crisis increase in mortgages puts banks more at risk related to sudden changes in the housing market. However, falling housing prices lead to a reduction in households' wealth and decreased demand, which in turn may lead to unemployment and unstable household income and therefore loan losses for banks. At the same time, as it was already mentioned, unemployment can be considered as a cause of decreasing demand and then contributes again to loan losses for banks.

Commercial property index is mainly connected to lending to the industries related to rental business and property management activities. A decrease in rental price leads to lower earnings and deterioration of collateral, as commercial property is usually the main collateral underlying enterprises' borrowing, especially for industries related to the rental market and commercial property management.

Capital buffer/equity-asset ratio

Adequate *capital buffers* provide a backup for loan losses because banks can deplete buffer capital before they reach a regulatory minimum of capital. Then the size of the buffer capital reflects how much loss the bank can absorb without necessary injections of new capital. A similar measure is the choice of the *equity-asset ratio*. These variables may have an ambiguous affect on loan losses. Banks may be willing to take higher credit risks if they hold larger equity capital and do not risk insolvency. Growing equity market and therefore stronger equity-asset ratio can create additional stimulus for risk-taking behaviour, while decreasing equity market accompanied by increased uncertainty and lower expectations lead to lower capital buffers and also increased risks due to the worsening of corporate accounts. Then, other things equal, banks may be less willing to take risks. At the same time, if bank managers value bank solvency and soundness quite low and prefer to keep low equity-assets ratio, they may also prefer higher and more volatile profits and may take higher risks.

Thus, the size of the equity-asset ratio allows us to incorporate the influence of bank buffer to withstand risk and shows bank willingness to take risk. A more general measure is a capital-asset ratio but it is less informative as its rise may also happen due to the increase in loss reserves.[22] Large banks usually have changes in capital-asset ratio due to the increase in loss reserves or decrease in assets, while we are more interested to track changes in equity.

Capital buffer safeguards against unexpected risks of the banks loan portfolio. These risks can be connected to the macroeconomic downturns, payment problems or bankruptcies of individual enterprises, increased lending to the corporate sector, concentration in particular industries, lower risk pricing and expansion to new customers. The latter factors are associated with intensified competition in banking. A bank with low capital, i.e. just above the minimum capital adequacy requirements, have high probability of being perceived as risky in the market, and therefore will have to borrow on worse terms and may experience liquidity problems.

Growth in loan portfolio

Rate of growth in loan portfolio reflects a rate of bank expansion in lending. High loan growth contributes to the reduction in capital adequacy, and therefore banks need solid profits to maintain funding and cannot sustain high loan growth for a long time. So lending is limited by the capital adequacy requirement when banks would like to raise new equity capital through new issues. Fast increase in lending may also cause higher loan losses through lower credit standards and larger increase in bad loans than in loans to creditworthy customers.[23] Moreover, lowering of credit standard compensated by lending margin may be followed by higher degree of moral hazard and adverse selection.

Non-performing loans

Non-performing loans are loans that have not been written off but are at least 90 days overdue, non-accruing or other problem loans with renegotiated terms. A change in the credit risk has an impact on the size of non-performing loans and non-performing loans net of loss provisions (net non-performing loans).

The size of non-performing loans reflects already defaulted (overdue) loans and can be different for banks with dissimilar lending specializations, and therefore reveals different information than aggregate bankruptcy probabilities. There is a time span between changes in credit risk and recorded problems with loans, as an enterprise with liquidity problems may not default on the loan if its shareholders agree to inject new capital. Banks may also undertake some loan restructuring, i.e. payment extensions, favourable change in terms of loan agreements, etc. Then loans are not considered non-performing. Aggregate bankruptcy

[22] Loss reserves are not used in the bank balances after 1995 and are excluded from equity in the data before 1995.

[23] However, Keeton (1999) argues that a relation between loan growth and losses does not occur only due to supply side which can be associated with softening of lending terms, i.e. lower interest rate, lower collateral requirements, lenient assessment of borrowers, etc. Changes in the demand and productivity can also cause an increase in lending when it comes along with the tightening of the credit standards.

probabilities and size of non-performing loans or a ratio of non-performing loans are only weakly correlated with coefficient of correlation around -0.02/0.12, while rates of loan losses are strongly related to the current level of non-performing loans.

The development pattern of the non-performing loans differs also from the change in the number of bankruptcies. Establishments and bankruptcies of small enterprises are quite common especially in some sectors of the economy. But some of these businesses are considered risky from the start and do not get ordinary loans.

Lending margin

Banks' financial results depend to a large extent on structure of lending and associated risks, and therefore loan pricing and credit risk measurement is an important component of banks' financial strategy. Lending margin reflects credit risk, goals for long-term profitability, administration costs and costs of funding. The size of risk included in the lending margin can also depend on the valuation of collateral because during the upward trend in the housing prices banks have lower risk of loan portfolio default. European financial reviews reflect a common tendency to a better risk management and greater importance of adequate risk pricing of loans. Increased lending margin and holdings of larger equity capital can lead to the same conclusion in Norway. Lending margin was reduced after the years of crisis only in 1994 and then after a short time was increased again in 1998.

In this study lending margin is calculated as a difference between bank's interest rate on loans and interest rate paid on the three month treasure bills, a proxy for the money market rate (i.e. marginal funding costs). Due to the varying banks' policies with respect to costs and risk pricing, lending margin is a more useful variable for explaining loan losses than macroeconomic changes in the interest rate, which have less direct effect on the ability of enterprises to serve their debt.

A decrease in banks' lending margin carries a possibility that banks' pricing policy is too mild and does not adequately reflect risks associated with corporate lending. In the conditions of intensified competition some banks review their pricing policy to win market shares. They can do this by reducing cost, by pricing risk lower and by decreasing their profits on loans. If this happens as a consequence of lower cost and better risk management then the bank can compete on the loan market maintaining its financial wealth, otherwise the risk of loan portfolio will markedly increase while earnings will deteriorate. Therefore, a decline in lending margins may increase banks' vulnerability to future losses on loans as risk may be priced inadequately.

Risk management/management quality

It is quite difficult to find an adequate proxy for the quality of banks credit policy. *Management quality* may be proxied by various profitability characteristics (i.e. the size of earnings before losses related to assets, return on equity) or cost effectiveness (i.e. total operating expenses related to average total assets). Profitability reflects bank's ability to generate revenue to cover incurred costs, pay dividends and retain profit. Banks may have increased profitability due to the increase in the rate of return or due to the change in the composition of assets and liabilities. But changes in return to assets, which are net of loss provisions, usually reflect changes in the size of the latter and thus may be misleading for our model.

Risk aversion

Lower risk aversion may cause banks to value profit possibilities more than possible costs of risk taking. Then unstable profits will cause much higher losses to the banks in case of macroeconomic shock, and this will be also aggravated by the influenced of these adverse shocks on financial situation of the banks' risky borrowers. Large variability in earnings and higher than average losses can serve as an indicator of risk-taking, i.e. banks with higher losses tend to have superior profits in the previous years and possibly charge higher interest on their loans to compensate for risk. As an indirect evidence of high-risk taking we can consider a loan to asset ratio, especially to risky industries or industries where higher interest rates are charged.

Risk diversification

A well-diversified bank may have lower risk as investments are spread over various industries and regions. If a bank provides loans to the industry or region with high bankruptcy probability it increases the bankruptcy probability of its total loan portfolio, while loans to the industries and regions with low bankruptcy probabilities have a mitigation effect. Specialising in a particular group of loans will carry higher risks and therefore an increase in the expected loss because of the higher probability of bankruptcy in this group. Moreover, large investment in a particular group reduces diversification in loan portfolio. A bank with low degree of diversification may still have comparable risk due to the higher expertise in particular industries. However, small low diversified banks may also have to accept higher risk due to the stronger competition.

A proper diversification of credit risk may lead to a much lower risk associated with loans. Savings banks have lower risk due to the higher share of mortgages, and thus they can and may be willing to decrease their credit standards and make loans with a higher default probability among corporate borrowers. Then they can profit from the possibility of charging higher interest rate to a variety of borrowers of lower credit class but at the same time incur costs of higher probability of bankruptcy of their borrowers. Large share of mortgages decreases the variability of possible loan losses and therefore makes banks more willing to engage in such policy as they benefit more than they lose.

At the same time, risk-weighted debt shows similar patterns with cyclical movements for most of the primary industries and counties in Norway.[24] Therefore, loan losses in banks are also expected to have some cyclical pattern with limited diversification opportunities across industry groups. Moreover, data availability constraints us by the assumption that banks have the same bankruptcy probabilities on loans inside a particular industry or region. Thus we should be aware that while possible diversification opportunities across major industries are limited, they are not taken into account at all within the industries and counties.

Market power

The degree of market power of a bank has implication for bank loan losses through its influence on the size of lending and deposit margins and also incentives to monitor borrowers.[25] At the same time, the degree of market power lowers incentives to take excessive risk through increased charter value of the bank and thus the size of losses in the case of failure due to the excessive risk-taking.[26]

[24] Calculation of the risk-weighed debt was done in Eklund et al (2001)

[25] Caminal and Matutes (2002).

[26] See a discussion in Perotti and Suarez (2002).

Competition in banking

Bank competition has a positive effect on the efficiency but it may also lead to an excessive risk-taking. A bank can expand its credit portfolio by underbidding its competitors or by accepting borrowers with lower creditworthiness. In the situation of intensified competition, in order to have compatible earnings banks may either try to compete by cost reduction or begin to expand aggressively and attract new clients that may highly increase their risk exposure. The latter contributes to the strategy of entering new industries and regions where banks do not have information advantage.

There was a sharp increase in the number of bank branches as a result of increased competition and larger freedom in new branch establishment. However, rapid expansion in new industries and geographical regions put banks' lending portfolios under higher risk than average in these industries and regions. Expanding banks possess limited information about customers from new market segments where they have little experience in specific conditions and particular characteristics of the borrowers. So they either should increase their screening and monitoring costs or tolerate higher risk and compensate it with larger lending margin. The latter was more apparent in expanding and optimistic economic conditions. However, this provided wider scope for unexpected risk, which together softened capital regulations[27] created higher fragility in the banking. The other side of the expansion into new sectors was a myopic and herding behaviour of bank managers. Steigum (1992) suggests that deficient accounting made it possible for them to show high profits at the first stages independent of the loan quality due to the large initial charges on loans apart from the interest rate. Herd behaviour is consistent with a strategy to show high profits and expand when other financial institutions are doing so, otherwise bank managers are punished for unsuccessful policy in the short-term. This they can trade off with long-term benefit of non-herd behaviour. But under some conditions, herding is a prevailing rational strategy for all agents and can be another cause of following financial fragility.

Lower risk pricing contributes to a decrease in lending margin. The size of lending and deposit margin, and spreads between banks can serve as indicators of the strength of competition. Narrowing difference between interest margins in different market segments indicates stronger competition both for new and existing customers.

Difference between large and small/medium size banks

Large banks have proven to have sound loan portfolio partly due to the better risk management strategies, higher possibilities for diversification and advantage in monitoring (cost reduction). Default costs are relatively higher for banks with small borrowers, as they have to administrate more bankruptcies with small repayment amounts. Moreover, the probability of borrowers' default may increase even more if higher interest rate will lead to moral hazard problems and cause firms to take larger risks. At the same time, small banks are more likely to deal with small businesses, are more flexible and have better possibilities in resolving conflicts of interest. According to Boyd and Runkle (1993) small banks, which operate in restricted markets, receive higher economic rents. However, risk increases due to the expansion to new industries, regions and customer from new market segments of which they have little information and experience. Therefore, variables reflecting changes in the industry/region

[27] Following Steigum (1992), capital requirements for Norwegian banks were reduced to 6.5 per cent in 1985 and then even further, when regulation allowed equity capital to be replaced by subordinated loan capital.

composition in banks' loan portfolio may reflect not only willingness to expand to new market segments because of risk-taking or stronger competition, but also the difference between large and small banks.

5. Background information and estimation methods

Separately aggregated data for loans to households and non-financial enterprises is used to explain corresponding loan losses. The essential component in the regression equation for non-financial enterprises is therefore risk-weighted debt $\sum_{i \in N} p_i L_i$, where L is amount of short and long-term debt of enterprises and p is bankruptcy probability for each enterprise from the set N of non-financial enterprises. Theoretically L should be a loss given default, as generally the bank loses not the whole amount of the loan after the borrower's bankruptcy. From empirical data we can conclude that only around 30 – 50 per cent of the loan can be restored in the case of bankruptcy, but more detailed data on all loans is not available. A simple regression of total loan losses on the risk-weighted debt and housing index as a collateral proxy produces statistically and economically significant results with a good explanatory power.

It is reasonable to assume that banks are heterogeneous from their external characteristics, as size, scope of operations, earnings, to internal characteristics such as client and investment policy, risk management, i.e. indicators on risk taken, tolerance to risk and following amount of buffer capital, competitive behaviour and costs. At the same time, it is even more interesting to look at the heterogeneity over time due to the known bank crisis in Norway in the beginning of 90-s, as we would like to have good explanatory variables, which can reflect variation in loan loss before, during and after the crisis. The aim of the analysis is thus to estimate how risk profile imposed on the bank by chosen loan portfolios can explain loan losses during the period 1988 – 2001 and especially during the banking crisis of the early 90-s.

In this study we use panel estimation methods, which have higher estimation ability of the heterogeneous data by utilising two sources of variation in the data. While cross-sectional data helps to explain some relations relying only on the heterogeneity between individuals at the given moment in time and time series capture variations over time, longitudinal data addresses both inter-individual and between-individual variation. Even quite short time series but moderate size cross-sectional data provide good possibilities for explaining variations in the data.

We have unbalanced characteristics of the dataset because of the bank mergers, closure of banks and their subsidiaries, and new bank establishments. We observe around 150 - 170 banks in the sample, and the largest fraction, near 70 per cent of the banks, is observed during all the years. However, around 16 per cent are observed only in the first or first two years and then were merged and stopped to submit financial information. In general, quite a small fraction of banks appeared or dropped out from the sample after the crisis, but in general we can observe mostly sample attrition because of the mergers. It is possible to argue about existence of the selection problem in this context, as banks that are taken over are mostly inefficient ones and possibly suffered losses in the previous periods. But it is only one side of the problem as this cannot be the only reason of mergers (i.e. banks can merge due to the possible cost savings and economies of scale after the merger) and also some banks are established during the sample period. So I will assume that appearance and dropping of banks from the sample is exogenous and is not dependent on the bank losses.

Variable		Mean	Std. Dev.	Min	Max	Observations
Bankruptcy prob.	overall	2.014316	.9619911	.1475461	8.172152	N = 1956
(per cent)	between		.6454108	.5982204	4.414628	n = 186
	within		.793504	-.4746066	6.862702	T-bar = 10.5161
Loan loss	overall	27957.18	213727.3	-802694.1	5117471	N = 1956
(mil NOK)	between		113285.9	-2653.735	872844.9	n = 186
	within		185283.9	-1468845	4451320	T-bar = 10.5161
Ratio loss-assets	overall	.0057483	.0167285	-.4049709	.3741996	N = 1956
	between		.0160808	-.0245717	.131978	n = 186
	within		.0153142	-.374651	.4045196	T-bar = 10.5161

Overall and within deviation is calculated for *N* bank-years of data. Between deviation is calculated over *n* banks. The average number of years a bank is observed is 10.5. For example, average bankruptcy probability is 2 per cent with standard deviation 0.962 per cent and it varies between min = 0.148 per cent and max= 8 per cent over the considered 13 years. Average risk indicators for each bank for 13 years have lower standard deviation of 0.645 per cent and lie in a smaller range between 0.598 and 4.4 per cent. Within number show deviation from each bank's average over time which also explains negative sign for the minimum, but we also need to deduct global means[28] and so they vary between -0.475 – 2.014 to 6.863 – 2.014. We also see that a deviation observed within banks over time is higher for risk indicators and loan losses but lower for the loan loss ratio than variation across banks. But we observe high variation in the data both between banks and over the years.

Econometric model

The analysis of the constructed longitudinal dataset is aimed to investigate whether calculated aggregate risk indicators for banks are significant and can explain, at least to some extent, bank loan losses. The following model is considered:

$$LA_{it} = \alpha_i + ABP_{it-1}\,\beta + M_t\xi + S_{it}\,\rho + \varepsilon_{it}\,, \qquad i \in 1{:}N,\ t \in 1{:}T \tag{1}$$

where N is number of banks, T is the number of periods equal to 12 and disturbances ε_{it} are identically and normally distributed with zero mean and constant variance σ^2. Variable LA is calculated as a ratio of loan losses to the total bank assets. This variable is more of interest than simply bank loan losses, as the variability of the loan losses can be huge not only due to the risk in lending but also due to the diversification effect related to the bank size and size of the loan portfolio, along with other factors. Variable ABP_{it-1} is a one period lagged aggregate bankruptcy probability, a risk indicator for a bank's portfolio of corporate loans. We use lagged values as mostly values realised in the previous period may influence losses of the current period.[29] Variable M_t stands for some of the macroeconomic variables (e.g. GDP, unemployment, and housing price index) and variable S_{it} stands for a vector of time- and bank-

[28] We can transform the model as follows:

$LA_{it} - \overline{LA}_{i\bullet} = (ABP_{it-1} - \overline{ABP}_{i\bullet})\beta + (M_t - \overline{M})\xi + (S_{it} - \overline{S}_{i\bullet})\rho + (\varepsilon_{it} - \overline{\varepsilon}_{i\bullet})$, where averages over years are calculated as: $\overline{LA}_{i\bullet} = \frac{\sum_t LA_{it}}{T}$. Estimated model have also global means added to each intraindividual difference. $LA_{it} - \overline{LA}_{i\bullet} + \overline{LA} = \alpha + (ABP_{it-1} - \overline{ABP}_{i\bullet} + \overline{ABP})\beta + (M_t - \overline{M})\xi + (S_{it} - \overline{S}_{i\bullet} + \overline{S})\rho + (\varepsilon_{it} - \overline{\varepsilon}_{i\bullet} + \overline{\eta}) + \overline{\varepsilon}$

[29] Presence of the lagged regressors makes it necessary to take into account bank mergers, which were especially widespread during the beginning of 90-s.

specific characteristics. Loan losses can take negative values because banks make reversals of previously made loss provisions if they overestimated their size and some of the breached contracts were repaid next period or they value given default happened to be higher than expected. Therefore, we do not use logarithmic form of the equation, which would be useful for log-normal distribution of positive values of loan losses. Due to data construction of variables are predetermined in the model and are assumed to be exogenous and uncorrelated with the disturbance term.

6. Estimation and hypothesis testing

Two different banks may invest in the same industry and region but have different investment results due to the diverse credit policies and different client base. A major shortcoming of the constructed aggregate bankruptcy probabilities is that we have to assume the same average credit risk for the banks that have loans to the same industries and region. However, loan losses dependent not only on the size of loans to riskier industries but also on the size of loans provided to more financial fragile enterprises. Therefore we have to use proxies that can help to distinguish banks with respect to their lending policies, i.e. quality of risk management, inclination to take risks and expansion into the new regions and industries, see discussion in section 4.

We conducted an estimation of the random effect model, for which individual specific effects η_i are correspondently assumed to be constant or randomly distributed. As α_i can be decomposed into a constant and individually variable part, we can rewrite the model as:

$$LA_{it} = \alpha + ABP_{it}\beta + M_t\xi + S_{it}\rho + \eta_i + \varepsilon_{it}, \quad i \in 1{:}N,\ t \in 1{:}T, \tag{2}$$

where $\eta_i + \varepsilon_{it}$ is a composite error term composed of the genuine disturbance and individual effect part, which is supposed to be randomly distributed and η_i to be drawn from the same probability distribution with IID $(0, \sigma^2_\alpha)$ and ε_{it} is IID $(0, \sigma^2)$ as before. We also make a strong assumption of independency of η_i, ε_{it} and explanatory variables. So we have non classical gross disturbance due to heteroskedasticity and autocorrelation through the variance of the individual random effect, and thus estimate the model by GLS.[30]

[30] Generalised least squares provides a weighted estimate of β using both within and between variation and assigning a smaller share to the 'between' one. This share is smaller when we have larger part of the gross disturbance variance due to the individual random effect. In our case we have almost 1/5 of gross disturbance variance due to the random individual effect.

Table 2: Random effect GLS regression

Dependent variable	Ratio loss/assets	Ratio loss/assets	Ratio loss provisions/assets
Aggregate bankruptcy probability (ABP)	0.002 *** (0.0006)	0.009 *** (0.002)	0.004 *** (0.0006)
Difference of the ratio of non-performing loans to assets	0.392 *** (0.019)	0.306 *** (0.014)	
Ratio of non-performing loans to assets			0.170 *** (0.007)
Unemployment		0.001 *** (0.0004)	-0.00005 (0.0001)
Ratio equity to assets		0.004 *** (0.001)	-0.003 *** (0.0004)
Share of risky loans		0.002 (0.003)	0.005 ** (0.003)
Interest rate $_{t-1}$	0.0005 *** (0.0002)		
Number of regions	0.0008 ** (0.0003)	0.0005 *** (0.0002)	0.00001 *** (0.000)
Dummy sb (if saving bank then 1)		0.011 *** (0.003)	0.0005 (0.0013)
Dummy sb*ABP		-0.007 *** (0.002)	-0.003 *** (0.0007)
Constant	-0.008 *** (0.002)	-0.017 *** (0.004)	-0.001 (0.001)
Breusch and Pagan LM test for RE[31]: Test: Var(u) = 0		chi2(1)= 18.33 Prob>chi2=0.000	
Hausman test: Ho difference in coefficients not systematic[32]		chi2(6)= 3.88 Prob>chi2=0.794	
R^2: within	0.279	0.250	0.329
between	0.671	0.436	0.651
overall	0.293	0.289	0.428

Macroeconomic variables are strongly correlated with coefficient of correlation -0.88 for unemployment and GDP, and 0.9 housing index and GDP. As macroeconomic variable we choose unemployment due to the above mentioned valuable properties for explaining loan losses. To reflect bank-specific variables we consider a ratio of non-performing loans to assets, a share of risky loans[33], interest rate on loans, a ratio of equity to assets and number of regions in bank loan portfolio.

The model provides economically and statistically significant results with the expected coefficient signs but not very high explanatory power. Due to the asymptotic properties of the

[31] Breusch-Pagan (1980) Lagrange multiplier test supports the idea of the random effect model. Statistics distributed as $\chi 2$ with one degree of freedom, under the null hypothesis of no random effects (i.e. zero variance of the individual specific part of the gross disturbance), is equal to 18.33 and hypothesis can be rejected.

[32] Assuming our correctly specified model and uncorrelation of ηi and RHS variables, we check that two models do not give statistically different results. Hausman's (1980) specification test checks the null hypothesis that difference in coefficients is not systematic and it cannot be rejected with p-value equal to 0.79. Difference between coefficients is statistically insignificant as null hypothesis cannot be rejected (probability of error is 79 per cent), and we can use random effects estimator for our model.

[33] Defined as a share of loans to non-financial firms with bankruptcy probabilities higher than three per cent to total loans in the bank's loan portfolio.

random effect estimator, Wald statistics confirm presence of significant regression on the 95 per cent significance level.

Both unemployment and non-performing loans are found to have a positive effect on bank loan losses. In addition ratio of equity to assets and share of risky loans also have positive influence reflecting adverse incentives arising from higher capital buffer. The number of regions has a statistically significant positive coefficient, suggesting that larger expansion increases risks and creates adverse effect for the loan losses. In addition to evaluating statistical significance of the sign of the coefficients it is useful to check the plausibility of the size of the obtained effects. Non-performing loans have highest effect on loan losses as an increase on 0.01 in the ratio of non-performing loans to assets with the average value of 0.023 leads to a 0.033 percentage points increase in the ratio of loan losses from 0.0057 to 0.0087 on average. Aggregate bankruptcy probabilities have less apparent but still quite large and statistically significant result. An increase on a 0.1 percent from the average value of 2 per cent leads to an increase from 0.0057 to 0.0066 in the ratio of loan losses to assets, which is around 15 per cent increase compared to the average value of this ratio.

A dynamic specification of the model was also estimated as losses in one period can be driven by the previous periods losses, which, for instance, can capture prevalence of banks' inefficient policy in assessment of the borrowers' credit risks or/and reflect influence of the banks' financial conditions on losses through the past performance.[34] Meaningful and statistically significant results for the aggregate bankruptcy probabilities are robust to the dynamic specification of the model.

[34] Here we have to deal with the endogeniety problem, as explanatory variables are correlated with the disturbance term (i.e. violation of the weak exogeniety).

Table 3. Dynamic model GMM estimation (robust to heteroskedasticity)

Dependent variable	Ratio loan losses/assets	Ratio loan losses/assets	Ratio provisions/assets
Ratio loan losses to assets $_{t-1}$	-0.524 *** (0.144)	-0.467 *** (0.142)	
Ratio loan losses provisions to assets $_{t-1}$			0.704 *** (0.087)
Aggregate bankruptcy probability (ABP)	0.002 ** (0.0008)	0.002 * (0.0009)	0.0002 (0.0006)
Aggregate bankruptcy probability (ABP_{t-1})	0.003 *** (0.0008)	0.004 *** (0.001)	0.0006 (0.0006)
Aggregate bankruptcy probability (ABP_{t-2})	-0.001 * (0.0006)		
Ratio non-performing loans to assets	0.303 * (0.178)	0.292 (0.187)	0.031 (0.013)
Equity/assets	0.022 (0.083)	0.027 (0.079)	0.033 *** (0.008)
Equity/ assets $_{t-1}$	0.475 ** (0.239)	0.482 * (0.247)	0.075 *** (0.027)
Unemployment		0.001 ** (0.0006)	0.0008 *** (0.0003)
Real loan interest rate (RIL)	0.003 *** (0.0009)	0.003 *** (0.0008)	0.0003 *** (0.000)
RIL $_{t-1}$	0.003 *** (0.0008)	0.002 *** (0.0006)	0.0002 * (0.000)
RIL $_{t-2}$	0.0009 *** (0.0003)		
Share of risky loans	0.006 (0.004)		
Share of risky loans $_{t-1}$	0.005 * (0.003)		
Number of regions	0.0005 * (0.0002)	0.0003 (0.0003)	-0.0001 (.0001)
Number of regions $_{t-1}$	0.0004 (0.0003)	0.0005 * (0.0003)	0.00007 (0.000)
Constant	-0.0007 *** (0.0003)	-0.0003 (0.0002)	0.0002 * (0.000)
A-B test: for zero 1-order aurocovariance in residuals	z = -2.28 Pr > z = 0.0228	z = -2.30 Pr > z = 0.0215	z = -2.03 Pr > z = 0.0423
A-B test: for zero 2-order aurocovariance in residuals	z = 0.41 Pr > z = 0.6821	z = 0.72 Pr > z = 0.4721	z = 0.21 Pr > z = 0.8302

7. Conclusions

We found that aggregate bankruptcy probability as a proxy for risks in lending can explain bank loan losses. This means that banks with higher bankruptcy probabilities of their loan portfolios tend to have higher loan losses if we control for the general macroeconomic conditions and bank specific factors. However for a given phase of the economic development, banks with more efficient credit risk management may be able to control risks more efficiently and reduce possible loan losses.
Agenda for the future research contains a possibility of testing of the following hypothesis:

1. Banks with higher level of management are likely to have lower loan losses for the same risk profile.
2. Banks that have better client base and have advantages in building superior client relationships will tend to have lower loan losses for the same direction of investment on the level of industry and region, i.e. the same aggregated bankruptcy probability for their loan portfolio.
3. Banks that are less risk averse may have riskier client base even if they choose the same direction of investment in terms of industries and regions as the other banks. (size of equity capital as a proxy).
4. Banks that have higher expectations regarding bail-out policies may take higher risks and allow higher loan losses (too-big-to-fail hypothesis).
5. Savings banks have larger investment in household sector and therefore can tolerate higher risks on their corporate part of the loan portfolio.
6. Try to capture possible shifts in the banks' lending policy when they enter markets of new regions and industries, which is not considered if we base our estimation only on general macroeconomic environment and information on total outstanding and problem loans.

Further improvement can be also done on finding better proxies for bank-specific characteristics in relation to the risk taking and risk management.

There is no widely accepted economic theory on banks' loan losses but it is interesting to incorporate a theoretic motivation that can be taken from both fields of economics and finance. Starting from 1970s financial economic theory provides a possibility to look on the bank as a productive firm, which maximises its profit transforming deposits (inputs) into loans (output). Decisions can be made on interest rate and amounts of inputs-output. Emphasis is also widely made on the asymmetric information issues in banking, in client relations and general decision-making. This can give a good basis for the econometric analysis in explaining bank decisions.

References

Beattie et al., (1995) *Banks and Bad Debts: Accounting for Loan Losses in International Banking*. Chichester: Wiley.

Becketti, S. and C. Morris, (1992) Are Bank Loans Still Special? *Economic Review*, Federal Reserve Bank of Kansas City, 71-84.

Berger A.N. and R. DeYoung, (1997) Problem Loans and Cost Efficiency in Commercial Banks, *Journal of Banking and Finance*, 21, 849-870.

Bernhardsen, E., (2001) A Model of Bankruptcy Prediction. Working Paper 2001/10, Norges Bank.

Boyd, J., P. Gomis, S. Kwak, B. Smith, (2001) A User's Guide to Banking Crises. Mimes. University of Minnesota.

Boyd J.H. and D.E. Runkle, (1993) Size and Performance of Banking Firm: Testing the Predictions of Theory. *Journal of Monetary Economics*, 31, 41-67.

Caminal, R. and C. Matutes, (2002) Market Power and Banking Failures. *International Journal of Industrial Organization*, 20, 1341-1361.

Chirinko, R. and G.D. Guill, (1991) A Framework for Assessing Credit Risk in Depository Institutions: Toward Regulatory Reform. *Journal of Banking and Finance,* 15, 785-804.

Diamond, D., (1991a) Monitoring and Reputation: the Choice Between Bank Loans and Directly Places Debt. *Journal of Political Economy*, 99, 689-721.

Eklund, T., K. Larsen and E. Bernhardsen (2001) Model for Analysing Credit Risk in the Enterprise Sector, The Central Bank of Norway, *Economic Bulletin,* 3, 99-106.

Fernandez S., J. Martinez, J. Saurina, (2000) Credit growth, problem loans and credit risk provisioning. The Central Bank of Spain. WP No.0018

Financial Stability Review, Bank of England, (2000-2002).

Frydl, E.J., (1999) The Length and Costs of Banking Crises. WP/99/30, IMF.

Herring, R., (1999) Credit Risk and Financial Instability. *Oxford Review of Economic Policy*, 15/3, 63-79

Hoggarth, G., R. Reis, and V. Saporta, (2002) Costs of Banking System Instability: Some Empirical Evidence. Journal of Banking and Finance, 26, 825-855.

Keeton W., (1999) Does Faster Loan Growth Lead to Higher Loan Losses? *Economic Review*, Federal Reserve Bank of Kansas City, 57-75.

Liu, C. and S. G. Ryan, (1995) The Effect of Bank Loan Portfolio Composition on the Market Reaction to and Anticipation of Loan Loss Provisions. *Journal of Accounting Research*, 33, 77-94.

Mailath, G. and L. Mester, (1994) A Positive Analysis of Bank Closure. *Journal of Financial Intermediation*, 3, 272-299.

Matutes C. and X.Vives, (2000) Imperfect Competition, Risk Taking and Regulation in Banking. *European Economic Review*, 44, 1-34

Mishkin F.S., (1991) Assymetric Information and Financial Crises: A Historic Perspective, R.G. Hubbard (eds), *Financial Markets and Financial Crises*, University of Chicago Press, 69-108.

Perotii, E.C and J.Suarez, (2002) Last Bank Standing: What Do I Gain if You Fail?. *European Economic Review*, 46, 1599-1622.

Salas, V. and J. Saurina, (2002a) Deregulation, Market Power and Risk Behaviour in Spanish banks. *European Economic Review*.

Saunders, A., (1999) *Credit Risk Measurement. New Approach to Value at Risk and Other Paradigms.* New York: Wiley.

Scholes, M.S., P.G. Wilson, and M.A.Wolfson, (1990) Tax Planning, Regulatory Capital Planning, and Financial Reporting Strategy for Commercial Banks. *Review of Financial Studies,* 3. 625-650.

Steigum, E. (2004) Financial Deregulation with a Fixed Exchange Rate: Lessons from Norway's Boom-Bust Cycle and Banking Crisis. In Trovald G. Moe, Jon A. Solheim and Bent Vale (eds), The Norwegian Banking Crisis. Norges Bank Occasional Paper no. 34. Forthcoming.

Apppendix: combining bank statistics on loans with statistics on non-financial enterprises

In order to estimate how aggregate risk indicators for banks' corporate loan portfolios can explain loan losses during the period 1988 – 2001, we construct aggregate bankruptcy probabilities for banks on the basis of individual bankruptcy probabilities for corporate borrowers. We combine two types of annual data: bank statistics on loans provided by the Norwegian savings and commercial banks and branches and subsidiaries of foreign banks in Norway, and statistics for non-financial enterprises. Statistics on loans and loan loss provisions on loans to different industries of the Norwegian limited liability enterprises sector is recorded in Report 50 and Report 60, and total loan losses for each bank are available in Report 20.[35] Statistics on enterprises is taken from the SEBRA-database containing financial and other relevant information on Norwegian limited liabilities enterprises, and their bankruptcy probabilities are calculated on the basis of this statistics using the SEBRA-model[36].

Data about individual corporate borrowers is not available in the bank statistics, while the SEBRA-database contains disaggregated information. Therefore, we combine data from the bank reports and the SEBRA-database using different levels of aggregation and different types of industry classification. To construct a risk measure for a bank, individual bankruptcy probabilities are aggregated within county and/or industry groups using industry classification from the bank statistics and then weighted by the volume of loans granted to each of these groups by this bank. In this respect, a formal correspondence pattern between two industry classifications is utilised, where possible; whereas some artificial relationship between them is suggested, where necessary. Besides, data are examined and selected according to the demands of the analysis, while data peculiarities and some possible negative consequences of them are discussed below.

Statistics on bank loans

A description of the data:

From the Report 60 we take annual data on aggregate volumes of loans provided by the Norwegian savings and commercial banks and branches and subsidiaries of foreign banks in Norway to the sector of limited liability enterprises classified by industry and county over the years 1988 – 2001. Data on loans provided by other financial enterprises and mortgage companies, which constitute almost 40 per cent of all observations (around 20 per cent in volume of loans), are available only from 1996 and are not included in the data set. The type of industry classification in the bank statistics (i.e. the number and contents of specified industries) varies with 32 industries before 1991, 33 industries up to 1996, 58 industries in 1996-1997 and 59 industries up to 2001. Since we use data classified by industry, an introduction of a more detailed classification in 1996 can explain the rise in the number of bank loans to more than 6,400 compared to around 4,200 observations in the previous years. Then it continues to increase to over 7,000 in the year 2001, except for a small decrease in 1999, but the number of active banks decreases slightly during these years.

According to this classification, we can combine data from the annual banks' balance sheets with the SEBRA-database along two dimensions: industry dimension and industry/county

[35] Official statistics (annual balance sheet data) for banks and investment banks.
[36] See Bernhardsen (2001) and Eklund, et al (2001) for a description of the model.

dimension (further referred to as groups). Classifications, valid in the bank statistics before and from 1996, have relationship patterns with Standard Industrial Classification that was valid before 1994 (SIC83) and that is valid from 1994 (SIC94) respectively. Since the latter is used throughout all the SEBRA-database (i.e. old codes were previously converted to the new ones), a correspondence between SIC83 and NACE (European standard underlying SIC94) is used to combine information from the bank statistics and the SEBRA-database. However, this correspondence is not one-to-one and an additional link to the old industrial codes in Report 60 is established (see section Linking of the bank statistics and the SEBRA-database).

Data on loans in the Report 60 are represented by a single post 'net loans' from 1988 to 1992 and by two separate posts 'gross loans' and 'loan loss provisions' from 1992. Calculation of net loans, i.e. gross loans minus loss provisions, for the years 1992 - 2001 gives negative results for some groups due to presence of loss provisions in groups with zero loans or provisions that are larger than loans. Negative net loans were obtained for around 340 industry/county/year groups and around 90 industry/year groups after dropping unclassified loans. This can be explained by data errors and discrepancy in the assignment of counties for loans and loss provisions in the bank statistics. Moreover, a preliminary examination of the data shows also that loss provisions made by a bank for a one monetary unit of its loans are only weakly correlated with borrowers' bankruptcy probabilities. This can partially be explained by different methods used by banks in calculation of their expected losses, but can also be due to the above-mentioned non-correspondence of groups of loans.

Therefore, data on loss provisions in Report 60 are considered non-reliable and subtraction of provisions from gross loans may introduce inaccuracy to the calculations. In order to avoid these discrepancies we will use gross loans for the period 1992 – 2001 and net loans for the period 1988 - 1992. However, this may have some side effects. First, it increases the size of banks' exposure to risk as gross loans include both expected and unexpected risk. While net loans expose banks' capital to a lower risk, i.e. only to unexpected risk, if we assume that banks incorporate in their loss provisions all expected credit risk on extended loans. However, this is probably not very important for our calculations of the risk indicator, because volumes of loans are used only as weights. Second, because of data limitations, data on gross loans cannot be used for the whole sample (i.e. a division between the posts for gross loans and loss provisions was introduced in 1992). This probably introduces a shift in the measure of risk exposure, as it is likely that credit risk models and distribution of loss provisions may vary between banks.

Transformations for the whole period 1988-2001:

- Consider only Norwegian savings and commercial banks, foreign banks and their branches in Norway (institutions with registration numbers larger than 1000 and less than 9800). (Note: Registration numbers below 1000 correspond to public lending institutions. Numbers above 9800 correspond to investment banks and mortgage companies, for which information is not available before 1996).
- Do not consider banks with registration numbers between 5201 and 6000, i.e. branches of DNB abroad, which are present in Report 50 and in Report 60.
- Consider only loans extended to the sectors 710, 719 and 719 (i.e. Norwegian limited liability enterprises, for which accounting data is included in the SEBRA-database).
- Control for negative numbers for loans (31 negative observations were deleted) and positive numbers for loan loss provisions (2 positive observations were deleted).
- Use only first two digits for the county codes (drop geographical codes).

- Drop observations with missing or zero industry and county codes (1221 observations were deleted).
- Sum observations over the currencies (Note: Loans are extended in two currencies, NOK and foreign currency converted to NOK).
- Join observations for counties 21 – 23 in county 21 (Svalbard) (Note: Counties 22 and 23 are not defined in the SEBRA-database). Join observations for county 2 (Akershus) and county 3 (Oslo) in county 3 (Oslo/Akershus) due to the geographical and economic interrelations of these counties.

Transformations for 1996-2001

- Drop industries: Mining of uranium (code 120) and Securities trading excluding financial institutions (code 650).
- (Note: Industry 120 appears in the bank statistics from 1996 but is not present in the SEBRA-database. Industry 650 appears in the bank statistics from 1997 and corresponds to the industry code 65238, which does not exist in the Standard Industrial Classification of the Statistics of Norway).
- Change numeration of industries for 1996, from 10 and 20 to 11 and 21 respectively, as the latter are used in the bank statistics from 1997.

Transformations for 1988-1995

- Sum observations over two lengths of loans. (Note: A division between post 30 (loans for less than 1 year) and post 60 (loans for more than 1 year) in the balance sheet was valid until 1995).
- Drop industries: Investment and holding companies (code 810) and Services related to banking and finance (code 820), which correspond to Securities trading excluding financial institutions (code 650) after 1996.

Statistics on enterprises (annual financial statements)

A description of the data:

The SEBRA-database contains information from annual financial statements of limited liability enterprises registered at the Norwegian register for business enterprises over the years 1988-2001. The SEBRA model is used to predict probabilities of bankruptcy for enterprises with book value of total assets exceeding 250,000-300,000 NOK. An observation is defined here as a record with financial and other relevant information submitted by an enterprise (referring to its unique identification number) available in the database for a particular year.

The data set contains 1,399,119 observations over 14 years. The number of enterprises submitting their financial records was constantly growing from 47,641 in 1988 to 137,201 in 2000 with a small decrease in 1994, but there is a large drop of more than 7 per cent in the last period of the data set, year 2001. At the same time, number of enterprises in different industries varies from just few to several thousands. Only few observations exists for the following industries: Transport via pipelines (codes 603 and 221), Private households with employed persons (code 950), Manufacture of tobacco products (code 160), Mining of coal and lignite (code 100), Public administration and defence, compulsory social security (code 750), Collection, purification and distribution of water (code 410), Recycling (code 370), Manufacture of coke, refined petroleum and nuclear fuel (code 230), Mining of metal ores (code 130), Tanning and dressing of leather (code 190), Manufacture of office machinery (code 300), and some others. This problem increases when enterprises are also divided between counties. On the other hand, most of the enterprises are registered in the counties Oslo/Akershus, Hordaland, Rogaland, Møre and Romsdal and Buskerud. Much less are registered in Finnmark, Aust-Agder, Sogn and Fjordane, Nord-Trondelag, and only few at Svalbard. Therefore, when we take into account both counties and industries the problem becomes more severe for the latter. This pattern is similar to the statistics on loans in Report 60. It may be explained by the relatively low level of activities in these counties and industries. Moreover, some enterprises provide accounting information not on a regular basis, and so information for some of the years is missing. This can clarify why some of the constructed industry/county/year groups, which have information on loans in the bank statistics, contain zero observations in the SEBRA-database. And what is more important, for these reasons, it is not possible to obtain aggregated bankruptcy probabilities for all relevant groups directly from the SEBRA-database. Below some examples are provided:

- If we look at the sample for 1996-2000, only few enterprises were registered in industry Private households with employed persons (code 950), and almost all of them are registered only in the counties[37] 3, 11, 15 and 19. However, bank granted loans to this industry also in the counties 1, 6, 12, 16, 17.
- There are no enterprises engaged in industry Inland water transport (code 612) or in industry Manufacture of other transport equipment (code 350), or in industry Research and development (code 730) in the central part of Norway, county 5. However, bank statistics contain information on loans to the enterprises which belong to these groups. The same is valid also for county 20 and industry 730, and for county 17 and industry Provision of energy and water (code 540).
- There are more records on loans extended to industry Recycling (code 370) in Troms than there are enterprises operating there.
- According to the SEBRA-database, all enterprises in Transport of oil via pipelines are registered only in Rogaland, county 11. According to the bank statistics, however, some of them belong to the counties 3, 10, 12, and 18. Also all enterprises in Manufacture of tobacco products belong to the county 3, while loans to this industry are recorded for county 10 over 1996-1999.
- The same is the case for enterprises in Mining of coal (code 100) and in Manufacture of coke, refined petroleum (code 230). There are only few of them in the database and they are registered in less than a half of the counties. However, according to the bank statistics, loans were given to these enterprises in other regions also.
- In most cases there are not more than 20 observations in each industry/county/year group that lack corresponding bankruptcy probabilities and hence could influence aggregation within the group. Industry Collection, purification and distribution of water (code 410), however, have many absent bankruptcy probabilities. Only a few enterprises in this industry are recorded in the counties near the west and south coast of Norway (10, 11, 14 and 15), no one is registered in the very north of Norway (19 and 20) and only a few in counties 16, 17 and 18. However, many banks grant loans to this industry in counties 10, 14, 16, 19 and 20, according to the Report 60. The same is true for industry Post and telecommunication (code 640) for the years 1988-1995. Most of it enterprises are registered in Oslo and Akershus and only a few in 6, 7, 9, 11, 12, 15, 16, while records in the bank statistics corresponds to different counties including 1, 4, 5, 8, 10, 17, 18, 19.
- There are more enterprises which belong to the northern counties (19, 20, 21) according to the bank statistics than it is referred to in the SEBRA-database for particular years. Bank statistics contain information on loans given to enterprises operating in Manufacture of food products and beverages (code 150) in 1997, Extraction of oil and natural gas (code 110),

[37] See the list of counties in the Appendix.

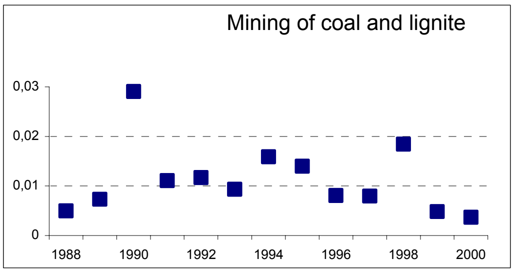
Manufacture of other engineering products (code 469), Business services in 1988 (code 840), Recreational, cultural and sporting activities (code 920) in 1996 and 2000 and Air transport (code 620) over 1996-2000 in county 21, enterprises in Manufacture of wearing apparel, dressing and dyeing of fur (code 180) in 1996, Manufacture of pulp, paper and paper products (code 210) over years 1996-1999, Manufacture of other electrical machinery and devices (code 310) in 1998 and Manufacture of medical, precision and optical instruments, watches and clocks (code 330) over 1996-2000 in the county 20, enterprises in Conversion of wood (code 320) over 1988-1994 in counties 19 and 20 and in Recycling (code 370) over 1996-2000 in county 19. For the corresponding years, there are no enterprises registered in the corresponding industries and counties, and submitting financial statements according to the SEBRA-database.

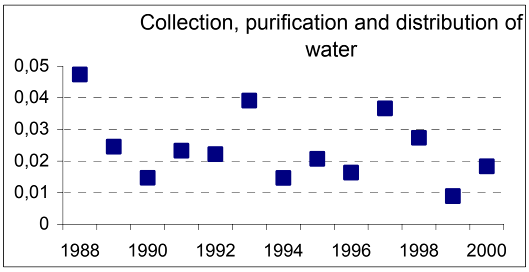
At the same time, a closer accord between the two statistics is obtained for the industry/year groups. For all groups constructed using the bank statistics, relevant aggregated bankruptcy probabilities can be calculated directly form the SEBRA-database after establishing a correspondence between industry codes in the two statistics. A minor exception is industry Public administration (code 750), to which loans are not registered in the bank statistics in 1996 and 1997, but some enterprises in this industry reported ‘debt in the financial institutions’ during the same years. However, a correct connection between groups made for enterprises with debt in financial institutions and groups of loans obtained from the banking statistics cannot be checked without full information on all financial institutions (Note: we consider only bank loans). Therefore, industry/year groups can be directly used for linking the bank and enterprise statistics, while industry/county/year groups require additional transformations.

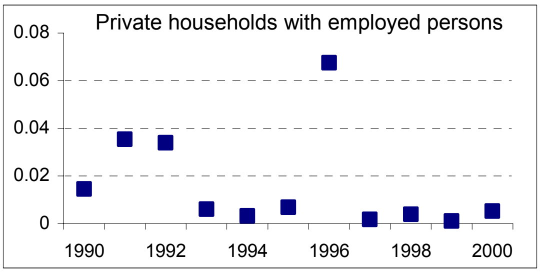
Information from financial statements also involves some problems. For example, in 17,835 cases out of all submitted financial statements for the period 1988 - 2000, assets do not match with the liabilities. For 296 observations assets are negative and are equal to the liabilities. Negative assets appear mainly due to the posts: cash, debtors, or investment in financial assets; however it is not the only source of problems. Negative liabilities show up mainly because of negative equity on the liability side of the balance sheet. Moreover, the same problem arises due to negative debt (4,452 observations), while it should be posted as assets, for example, as cash or deposits. Negative debt arises mainly due to bank overdrafts (short-term loans in financial institutions), but long-term negative liabilities are also present in the data. Negative operating revenue appears 625 times in the sample. However, it is not possible to find a concrete source of mistakes and to correct the figures without having more specific information about enterprises than accounting data contained in the SEBRA-database. In some cases, values of aggregated posts are not equal to the sum of the detailed posts. For example, the value of total assets cannot be obtained from the sum of fixed and current assets, and values of total liabilities cannot be obtained from the sum of equity and liabilities. However, using the calculated instead of the given numbers for total assets and total liabilities leads to a discrepancy between assets and liabilities in a large number of cases.

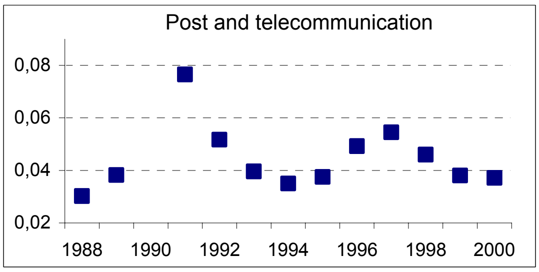
Another property of the data is that enterprises with recorded negative or zero assets have extremely low predicted probabilities of bankruptcy, almost all belonging to the range with $p < 0.01$, where p is the probability of bankruptcy. Those with negative operating revenue have more reasonable values of risk measures but most of them still belong to the range with lowest probabilities, i.e. $p < 0.01$. Therefore, bankruptcy probabilities estimated on these records are not trustworthy. Estimation of bankruptcy probabilities on the sample of enterprises with book value of total assets larger than 300,000 NOK still includes some records with negative revenues. The model was re-estimated setting them equal to zero; however, it did not significantly affect the results as these records constitute a small share of total observations. However, records with negative asset values should be excluded as predicted bankruptcy probabilities are not reliable and will introduce errors into the aggregated values.

Average bankruptcy probabilities in different industries show diverse patterns over the years 1988-2000. Most of them are decreasing, while some have slightly rising tendency over the last years, e.g. Forestry (code 021), Manufacture of pulp and paper (code 210), Manufacture of coke, refined petroleum products and nuclear fuel (code 230), Transport via pipelines (code 603) and Manufacture of tobacco products (code 160). Some had this tendency only till the middle of the sample period, with a following decrease after 1994 for Mining of coal and lignite (code 100) and after 1996 for Recycling (code 370). However, data on these industries, except 021 and 210 contains quite small number of observations, and results may be influenced by extreme values. Implausible values of the average bankruptcy probabilities are also found in the following industries: Mining of coal and lignite, Collection, purification and distribution of water, Private households with employed persons, Post and telecommunication.









High average bankruptcy probabilities with large deviations, i.e. mean value larger than 0.036 and standard deviation larger than 0.065, which corresponds to the upper 25 per cent, are found in many industries especially during 1990-1993. Values in the upper 20-10 per cent correspond are found in the following industries: Agriculture and forestry (code 111) in 1990-1992, Raising of fish (code 051) in 1988-1993, Fishery (code 052) in 1988-1993, Publishing, printing and reproduction of recorded media (code 220) in 1988-1993, Manufacture of office machinery and computers (code 300) in 1996-1999, Hotels and restaurants (code 550) in the whole sample 1988-2000, Construction (code 450) in 1989-1993, also industries 461, 462, 469 and 490 in 1990-1992, industry Wholesale and commercial agency (code 610), Trade retailing (code 621) in 1990-1991, Post and telecommunication (code 640) in 1990-1991 and 1996-1998, Recreation, cultural and sporting activities (code 920) in 1996-1999, Other service activities (code 930) in 1996-2000. These industries traditionally have high uncertainty in their activities; this is particularly true for hotel, restaurants, recreation and service activities. However, industry Real estate activities (code 700), which is also considered risky, show quite stable and low values of bankruptcy probabilities throughout the sample period with mean values decreasing from 0.013-0.016 to 0.007.

Data transformations:

- Drop observations with missing and zero industry code or zero county number (for years 1988-2001, 8124 and 262 respectively observations were deleted).
- Drop observations with industry codes that do not correspond with aggregate codes in the bank statistics:
  - Codes 65000 - 68000 (financial operations and insurance), (for years 1988-2001, 49793 observations were deleted);
  - Code 99000 (international organisations), (for years 1988-2001, 1450 observations were deleted);
- Drop observations with industry codes 75000 – 76000 before year 1996, because they do not correspond to the aggregate industry codes from the old classification (aggregate industry 190 according to the new classification).
- Drop observations with industry codes that do not correspond with SIC94 and cannot logically be added to one of the existing groups:
  - Code 38399 (contains records on one enterprise with average assets around 1.200.000 NOK, average revenue and debt around 800.000 NOK for the years 1992-1997, the highest value of debt was 2.146.000 NOK in 1995), (for years 1988-2001, 6 observations were deleted).
  - Code 83299 (contains records on one enterprise with average assets around 170.000 NOK, average revenue around 650.000 NOK for the years 1994-1997, the highest value of debt was 13.000 NOK in 1996, and on one enterprise with average assets around 1.200.000 NOK, average revenue around 2.000.000 NOK and average debt around 280.000 NOK over the years 1990-1996, the highest value of debt was 988.000 NOK in 1994), (for years 1988-2001, 11 observations were deleted).
  - Code 88888 (contains records on one enterprise with 1.845.000 NOK assets, zero revenue and 1.674.000 NOK debt in 2000), (for years 1988-2001, 2 observations were deleted).
- Classify observations with industry codes that do not correspond with SIC94 due to higher precision level and assign them codes from upper designation correspondent to SIC94. For the years 1988-2000:
  - 53 observations with code 01222 are added to the code 01220. Added records constitute 62 % of total observations (01222 and 01220 together), and 67 % in terms of extended loans;
  - 990 observations with code 01411 are added to the code 01410. Added records constitute 73% of total observations and 82% in terms of extended loans;
  - 9 observations with code 11111 are added to the code 11100. Added records constitute 0.9% of total observations and almost 0% in terms of extended loans;
  - 266 observations with code 20511 are added to the code 20510. Added records constitute 36% of total observations and 57% in terms of extended loans;
  - 3677 observations with code 28751 are added to the code 28750. Added records constitute 58% of total observations and 69% in terms of extended loans;
  - 1 observation with code 29012 are added to the code 28750. Added records constitute almost 0% in terms of observations and extended loans;
  - 173 observations with code 31630 are added to the code 31620. Added records constitute 84% of total observations and 95% in terms of extended loans;
  - 454 and 40 observations with codes 45001 and 45002 respectively are added to the code 45000. Added records constitute 51% of total observations and 6% in terms of extended loans;
  - 7902 observations with code 45111 are added to the code 45110. Added records constitute 81% of total observations and 86% in terms of extended loans;
  - 89 and 2173 observations with codes 45251 and 45252 are added to the code 45250. Added records constitute 47% of total observations and 23% in terms of extended loans;
  - 54 observations with code 51331 are added to the code 51330. Added records constitute 24% of total observations and 1.3% in terms of extended loans;

- 400 observations with code 51411 are added to the code 51410. Added records constitute 24% of total observations and 6% in terms of extended loans;
- 1256 and 1 observation with codes 51435 and 51439 respectively are added to the code 51430. Added records constitute 55% of total observations and 43% in terms of extended loans;
- 307 observations with code 51443 are added to the code 51440. Added records constitute 95% of total observations and 99% in terms of extended loans;
- 33 observations with code 51451 are added to the code 51450. Added records constitute 2.5% of total observations and 0.2% in terms of extended loans;
- 193 observations with code 51461 are added to the code 51460. Added records constitute 7% of total observations and 0.6% in terms of extended loans;
- 83 observations with code 51521 are added to the code 51520. Added records constitute 5% of total observations and 6% in terms of extended loans;
- 73 and 2 observations with codes 51552 and 51551 respectively are added to the code 51550. Added records constitute 4% of total observations and 1% in terms of extended loans;
- 70 and 391 observations with codes 51643 and 51641 respectively are added to the code 51640. Added records constitute 11% of total observations and 26% in terms of extended loans;
- 377 total observations with code 51701-51703 are added to the code 51700. Added records constitute 8.5% of total observations and 20% in terms of extended loans;
- 1127 total observations with code 51810, 51820, 51830, 51840, 51850, 51860, 51860, 51870, 51880 are added to the code 51000. Added records constitute 13% of total observations and 27% in terms of extended loans;
- 154 observations with code 52101 are added to the code 52100. Added records constitute 21.5% of total observations and 37% in terms of extended loans;
- 13812 total observations with code 52111-52113 are added to the code 52110. Added records constitute 47% of total observations and 62% in terms of extended loans;
- 28 observations with code 52221 are added to the code 512220. Added records constitute 2.5% of total observations and 1% in terms of extended loans;
- 328 observations with code 52601 are added to the code 52600. Added records constitute 97% of total observations and 99.5% in terms of extended loans;
- 154 observations with code 52741 are added to the code 52740. Added records constitute 61% of total observations and 80% in terms of extended loans;
- 337 observations with code 55401 are added to the code 55400. Added records constitute 24% of total observations and 29% in terms of extended loans;
- 8165 observations with code 61001 are added to the code 61000. Added records constitute 89% of total observations and 91% in terms of extended loans;
- 2 observations with code 63202 are added to the code 63200. Added records constitute 2% of total observations and almost 0% in terms of extended loans;
- 107 observations with code 64201 are added to the code 64200. Added records constitute 7.5% of total observations and almost 0% in terms of extended loans;
- 2 observations with code 63202 are added to the code 63200. Added records constitute 2% of total observations and almost 0% in terms of extended loans;
- 363 and 19 observations with codes 71402 and 71401 respectively are added to the code 71400. Added records constitute 18% of total observations and 10% in terms of extended loans;
- 3 observations with code 71911 are added to the code 71000. Added records constitute 2.6% of total observations and almost 0% in terms of extended loans;
- 245 observations with code 72301 are added to the code 72300. Added records constitute 7.5% of total observations and 1% in terms of extended loans;
- 13395 total observations with code 74401-74409 are added to the code 74400. Added records constitute 71% of total observations and 66% in terms of extended loans;
- 300 observations with code 74601 are added to the code 74600. Added records constitute 20% of total observations and 59% in terms of extended loans;
- 61 observations with code 74811 are added to the code 74810. Added records constitute 2% of total observations and 10% in terms of extended loans;
- 5147 and 350 observations with codes 74832 and 74831 respectively are added to the code 74830. Added records constitute 96% of total observations and 99% in terms of extended loans;

- 369 observations with code 74841 are added to the code 74840. Added records constitute 1.6% of total observations and 8% in terms of extended loans;
- 61 observations with code 74811 are added to the code 74810. Added records constitute 2% of total observations and 10% in terms of extended loans;
- 54 and 37 observations with codes 91331 and 91332 respectively are added to the code 91330. Added records constitute 3% of total observations and 3.5% in terms of extended loans;
- 513 observations with code 93012 are added to the code 93010. Added records constitute 22% of total observations and 4% in terms of extended loans;
- 5516 total observations with codes 93021-93024 are added to the code 93020. Added records constitute 97% of total observations and 96.5% in terms of extended loans;
- 1445 and 759 observations with codes 93041 and 93042 respectively are added to the code 93040. Added records constitute 88% of total observations and 95% in terms of extended loans;

In some cases added observations constitute a large part of the newly obtained groups both in terms of the number of observations and amount of loans. However, the SEBRA-database contains many enterprises, with industry codes not included in SIC94, that cannot be omitted.

Aggregation of the bankruptcy probabilities by industry and county

Using industry and industry/county dimension, common for the banks' Reports and the SEBRA-database, we aggregate bankruptcy probabilities for each enterprise from the SEBRA-model by industry or by industry/county. Volumes of debt in the financial institutions or the levels of activities, represented, for example, by total assets or operating revenues, are proposed as weights in aggregation. It is reasonable to concentrate on the enterprises with non-zero post 'debt in financial institutions' since only these enterprises will inflict a loss for the bank in the event of bankruptcy. However, aggregation using only debt as weights may cause some biases because the SEBRA-database contains post 'debt in financial institutions', which has a wider meaning than debt in banks. Therefore, some of the selected enterprises will still be irrelevant for the calculation of the bank's risk on the loan portfolio. Moreover, banks have also loans to some industry/year groups that are not reflected in the SEBRA-database as groups of enterprises with debt in financial institutions'. For example, the SEBRA-database does not contain enterprises with debt in financial institutions in industry Transport via pipelines (code 603) in 1997 and in industry Private households with employed persons (code 950) in 1996, 1998 and 2000. Enterprises with debt, which belong to a particular industry/county/year group in the SEBRA-database, do not always correspond to those enterprises that banks have actually given loans in this group, also due to the discrepancies in the county classification.

However, due to the discussed shortcomings of the data it is also problematic to choose one of the possible activity measures, i.e. total assets or operating revenue. Both of them include negative observations, also for enterprises with bank loans. Therefore, we use the level of debt in the financial institutions as weights in aggregation.[38]

[38] A composite measure taking into account both characteristics can also be relevant because of the high correlation between them: corr(total assets, debt) = 0.9, corr(revenue, debt) = 0.7, corr(total assets, revenue) = 0.8. For example, we can assign half weight to the debt size and half weight to the measure of the activity size, or even use both activity measures as enterprises with negative total assets may have operating revenue and vice versa.

Linking of the bank statistics and the SEBRA-database

Assigning new aggregate industry codes

Assignment of the aggregate industry codes, valid in the bank statistics from 1996, to enterprises in the SEBRA-database is made for the years 1996 - 2001 according to the relationship pattern between SIC94 and aggregate codes in the Report 60. Observations from the SEBRA-database with industry codes that do not have a direct correspondence with aggregate codes are subdivided as follows:

- 143 observations from industry Fishing, operation of fish hatcheries and fish farms (code 5000) are randomly equally divided between the aggregate industries Fishing and operation of fish hatcheries (code 051) and Fish farms (code 052). These randomly added observations constitute 1.48 and 0.63 per cent respectively of the total number of observations in these industries.
- All 2071 observations from industry Land transport, transport via pipelines (code 60000) are included in the aggregate industry Land transport (code 601), and constitute 7.55 per cent of the total number of observations in this industry. (Note: Industry Transport via pipelines (code 603) has only few enterprises with a total of 33 observations for the whole sample period.)
- 1046 observations from industry Water transport (code 61000) are randomly divided between the aggregate industries Foreign water transport (code 611) and Inland water transport (code 612). These randomly added observations constitute around 2 and 19 per cent respectively of the observations in these industries.

Aggregate industries with codes 051 and 052 are joined before 1991, because during this period industry 053 contains information on both 051 and 052. It is also suggested to exclude industry Private households with employed persons (code 950), which does not contain sufficient information. It has around 2-5 observations each year, of which in general one enterprise has debt. The total sum of debt for this industry from 1988 to 2000 is 619,000 NOK, total revenue is 106,139,000 NOK.

Assigning old aggregate industry codes

The SEBRA-database contains only industry codes consistent with SIC94 becasue they were previously converted from SIC83 for all enterprises, while bank statistics uses old aggregate classification of industries in Reports 60 up to 1996. Therefore, for the data before 1996, we should assign old aggregate codes to enterprises using relationship patterns between old aggregate codes and SIC83, and between SIC83 and SIC94. This correspondence is not one-to-one, i.e. a particular industry code in SIC94 can correspond to several different codes in SIC83 due to the different types of the classification applied, and hence to several aggregate codes. Since we do not possess detail information about sphere of activities of the individual enterprises, distribution of some observations can be only done randomly. Therefore, some groups of observations with assigned old aggregate codes contain an arbitrary part. How these random observations may influence our data set is discussed below.

The following procedure is applied when defining the old and new aggregate codes:

1. Each observation with industry code that has a one-to-one correspondence with SIC83 is attributed to a correspondent aggregate industry;
2. Observations with industry code that corresponds to more than one industry in SIC83 are randomly divided between relevant aggregate industries;

3. Observations that do not have detailed industry codes consistent with the relationship pattern between SIC94 and aggregate codes are randomly divided between corresponding aggregate industries. Consequently, they will increase the number of randomly distributed observations.

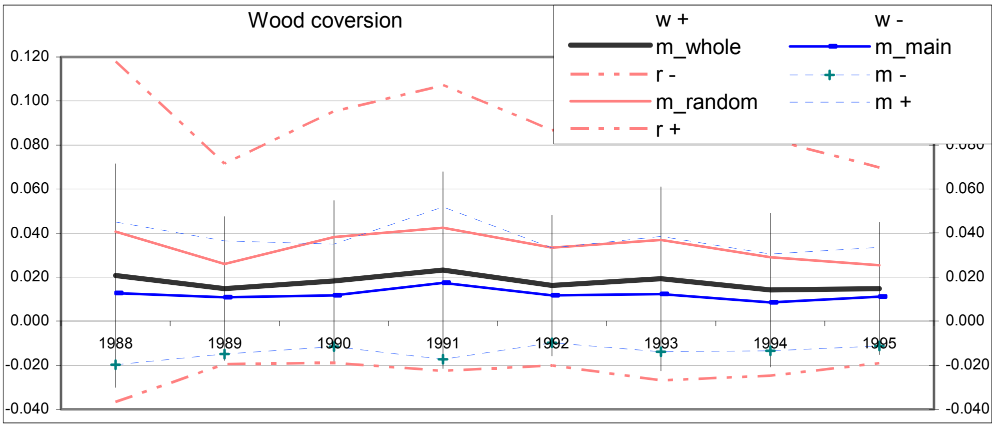
After the first examination, it turned out that some of the industries contain a very large random part, i.e. close to 100%. To avoid this, industries Transport and storage (code 911) and Foreign water transport (code 711) were united under the one industry 911 due to the similarities in their main activities. Aggregate industries Extraction of oil and gas (code 211), Financial operations relevant to extraction of oil and gas (code 231) and Drilling for oil on contact base (code 721) were joined in one industry 211 because they all have activities relating to oil and gas.

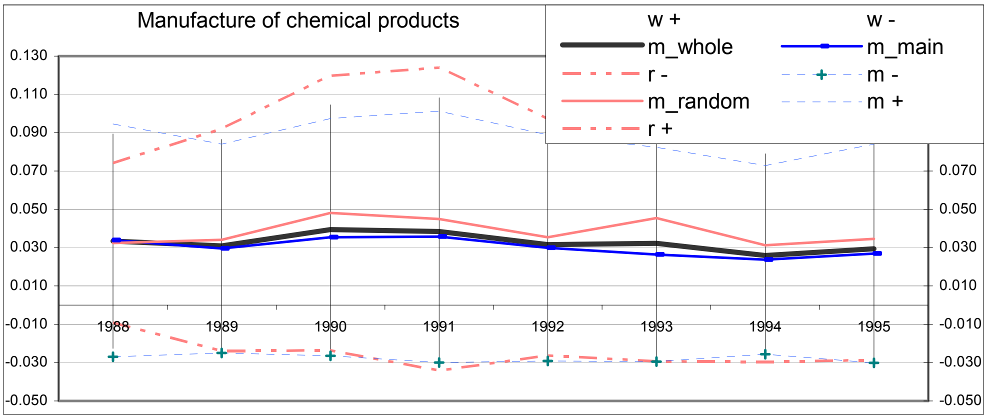
Then following old aggregate tree-digit industry codes were assigned to the observations with consistent with SIC94 five-digit numbers, using the correspondence between SIC94 and SIC83, and between SIC83 and aggregate codes:

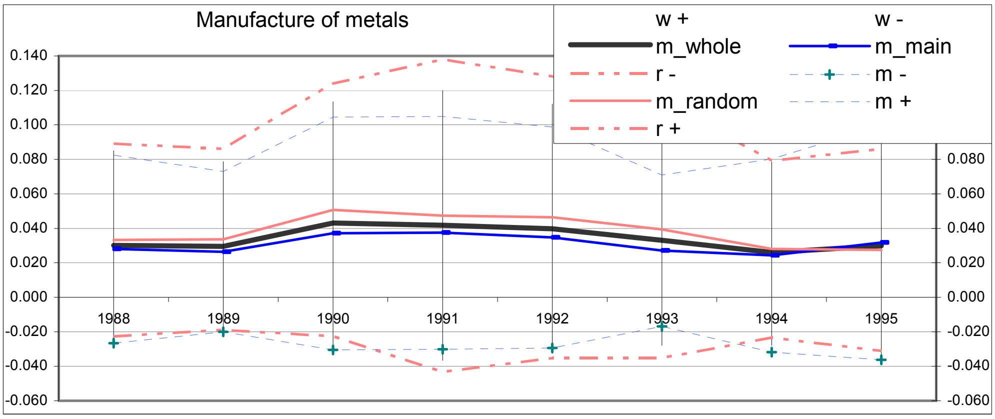
- >=01000&<01400, 01420, >=02000&<02020 correspond to 111;
- A half of observations with >=01400&<01420, 01500, 02020 was randomly selected for 111;
- >=05010&<05020 correspond to 051;
- A half of observations with 01500, 05000 was randomly selected for 051;
- >=05020&<10000 correspond to 052;
- A half of observations with 5000 was randomly selected for 052;
- >=11000&<13000 correspond to 211;
- A half of observations with 74203 was randomly selected for 211;
- 60300 corresponds to 221;
- >=10000&<11000, >=13000&<15000 correspond to 311;
- >=21000&<21230, >=21240&<22000 correspond to 321;
- A half of observations with 20200 and 21230 was randomly selected for 321;
- >=23000&<24139, >=24140&<25130, >=25200&<25240 correspond to 331;
- A half of observations with 24139, 26820, 29600, 35116, 35120, 36400, one third of observations with >=36100&<36200, 19300, 36500 and one fourth of 36630 were randomly selected for 331;
- Also 80 per cent of the observations in 25130 and 25240 was randomly selected for 331 (Note: 20% of observations in the industries 25130 and 25240 was allocated to 420 and the rest 80% to 331 because of the small extend of the textile industry);
- >=27000&<27210, >=27300&<27340, >=27350&<28000, 28400, 28500, 28520, 29140 correspond to 341;
- A half of observations with 24139, 28000, 29229, 34300, 27210, 27220 was randomly selected for 341;
- A one third of observations with 28700, 28750, 28751, 29200 and 29220 was randomly selected for 341;
- >=15000&<17000 correspond to 411;
- A half of observations with 55500 and 55520 was randomly selected for 411;
- >=17000&<19300 correspond to 420;
- A half of observations with 20520, one third of observations with 19300, one fourth of observations with 36630 and one fifth with were randomly selected for 420;
- >=20000&<20200 and >=20300&<20520 correspond to 430;
- A half of observations with 20200, 20520, 36000, 36150, 32300 was randomly selected for 430;
- A one third of observations with 19300, 31500, >=36100&<36150 was randomly selected for 430;
- >=22000&<22140 and >=22150&<22300 correspond to 440;
- >=26000&<26820 correspond to 451;
- A half of observations with 26820 and a one third of observations with 31500 were randomly selected for 451;

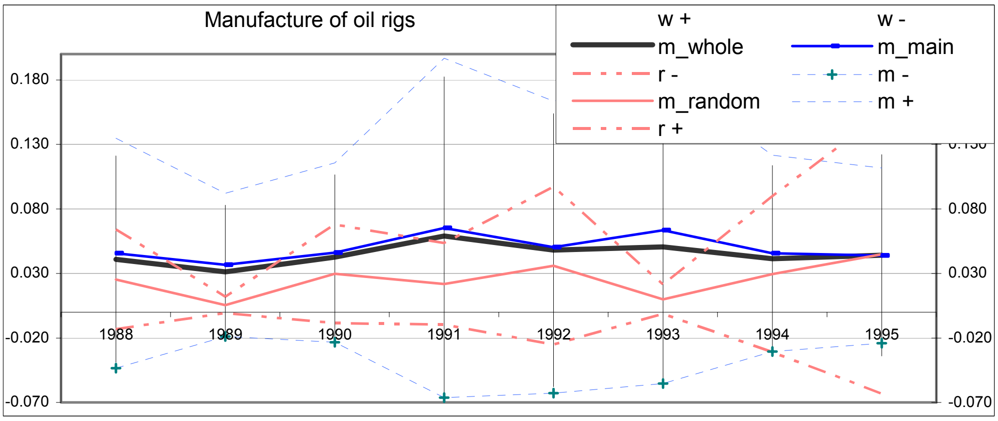
- >=35114& <35116 correspond to 461;
- A half of observations with 45212 was randomly selected for 461;
- 29111, >=35100&<35114 and 35117 correspond to 462;
- A half of observations with 29110, 29221, 35000, 29200, 35116, 35120 and a one third with 29220 was randomly selected for 462;
- 27340, >=28100&<28400, 28510, >=28600&<28700, >=28710&<28750, >=29000&<29110, >=28119&<29140, 28210, >=28230&<29600, >=28710&<31500, >=31600&<32300, 33000, >=32200&<34300, >=35201&fn<36000 correspond to 469;
- A half of observations with 27210, 27220, 45300, 45310, 45330, 51570, 72500, 28000, 29110, 29221, 29229, 29600, 32300, 33100, 34300, 35000 , a one third of observations with 28700, 28750, 28751, 29200, 29220, 31500, >=36100&<36150, 36500 and one fourth with 36630 were randomly selected for 469;
- >=36200&<36400 and >=36600&<36630 correspond to 490;
- A half of observations with 21230, 36000, 36400, one third with 28700, 28750, 28751, 36500 na one fourth with 36630 were randomly selected for 490;
- >=45000&<45120, >=45200&<45212, >=45220&<45300, 45320, >=45340&<50100 correspond to 450;
- A half of observations with >=1400&<1420, 2020, 45212, 45300, 45310, 45330 was randomly selected for 450;
- >=40000&<45000 correspond to 540;
- >=37000&<40000, 50101, 50301, 50401, 52464, 52485, >=51000&<51570, >=51600&<52000 correspond to 610;
- A half of observations with 52460, 52461, 52469, 50100, 50300, 51570 and a one third with 50400 were randomly selected for 610;
- 50102, 50302, 50402, 50500, >=52000&<52460, 52462, 52463, >=52470&<52485, >=52486&<52700 correspond to 621;
- A half of observations with 50100, 50300, 50200, 52460, 52461, 52469, a one third with 50400, >=71400&<72000 and a one fourth with 71000 were randomly selected for 621;
- >=55000&<55500, 55510 correspond to 550;
- A half of observations with 55500 and 55520 was randomly selected for 550;
- >=70000&<71000 correspond to 700;
- 22330, 45120, 71220, >=71300&<71400, >=72000&<72500, 72600, >=74000&<74203, 74209, >=74400&<74700, 74800, >=74820&<75000 correspond to 840;
- A half of observations with 22300, 92400, 71200, 71230, 72500, 74203, a one third with 74300 and a one fourth with 71000 were randomly selected for 840;
- >=60000&<60300, >=61000&<64100, 71100, 71210 correspond to 911;
- A half of observations with 64100, 64120, 71200, 71230, a one third with 74300 and a one fourth with 71000 were randomly selected for 911;
- 64110, 64200 and 64201 correspond to 640;
- A half of observations with 64100 and 64120 was randomly selected for 640;
- 22140, 22310, 22320, 50403, >=52700&<55000, >=73000&<74000, 74700, 74810, 74811, >=80000&<92400, >=92500&<99000 correspond to 931;
- A half of observations with 22300, 92400, 33100, 50200, >=71400&<72000, a one third with 50400, 74300 and a one fourth with 71000 were randomly selected for 640;

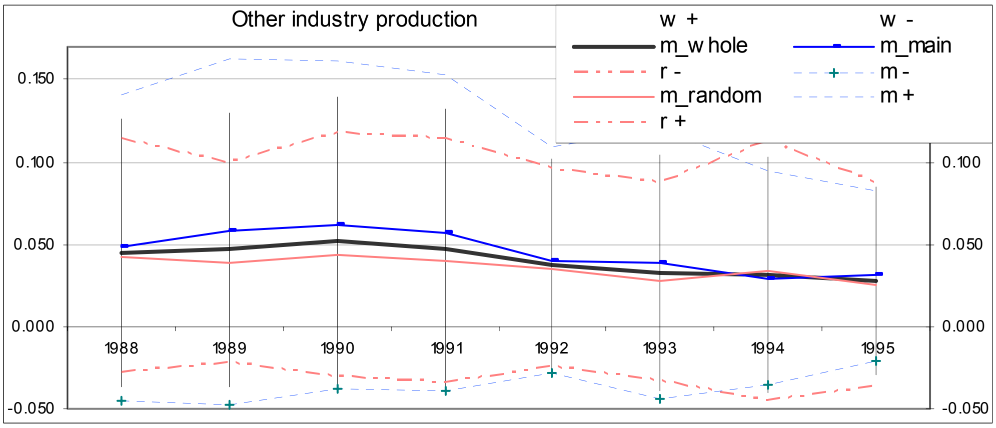
We analyse also the effect of the randomly added observations for the distribution of the bankruptcy probabilities *(p)* for the aggregate industries. The following graphs show the means and standard deviations for *p's* before and after the randomly selected observations were added, as well as for these random observations themselves. Variables m_random, m_main and m_whole are respectively mean values of the bankruptcy probabilities in the arbitrary added part, and mean values before and after this part is added over years. Variables r-/r+, m-/m+ and w-/w+ show respectively standard deviations in bankruptcy probabilities in the arbitrary added part, before and after this part is added.

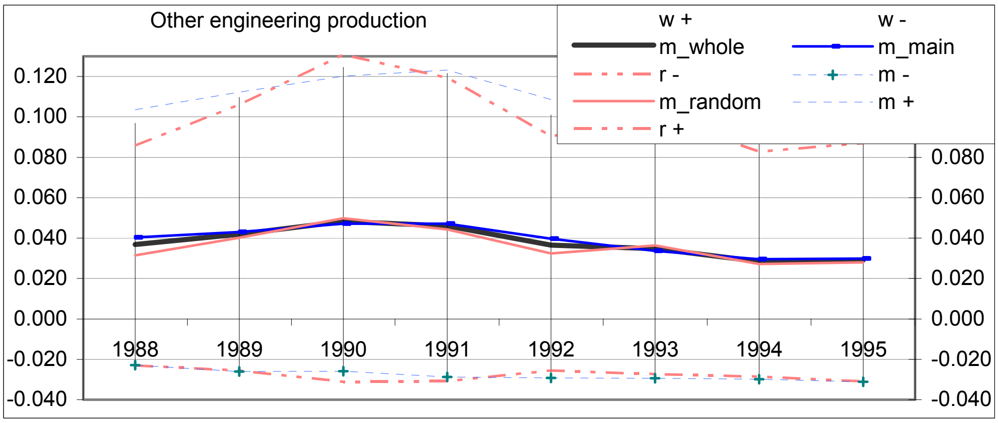

Wood coversion
w +
m_whole
r -
m_random
r +
w -
m_main
m -
m +
0.120
0.100
0.080
0.060
0.040
0.020
0.000
-0.020
-0.040
1988
1989
1990
1991
1992
1993
1994
1995

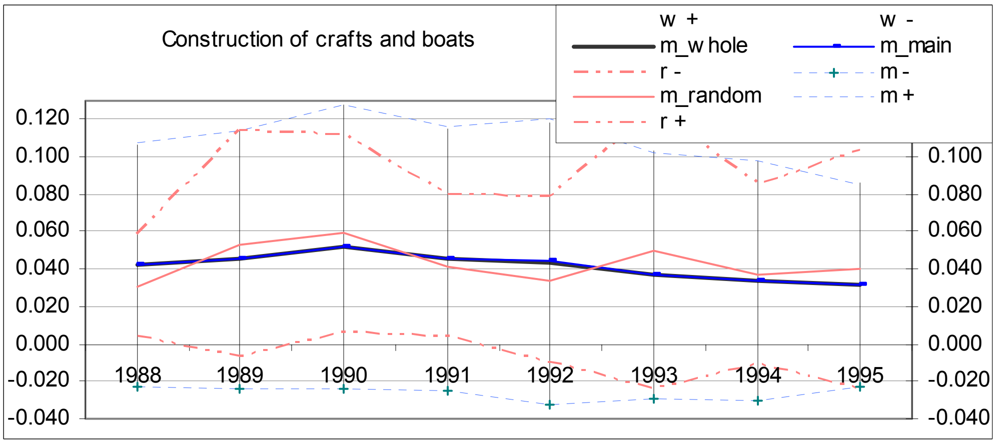

Manufacture of chemical products
w +
m_whole
r -
m_random
r +
w -
m_main
m -
m +
0.130
0.110
0.090
0.070
0.050
0.030
0.010
-0.010
-0.030
-0.050
1988
1989
1990
1991
1992
1993
1994
1995

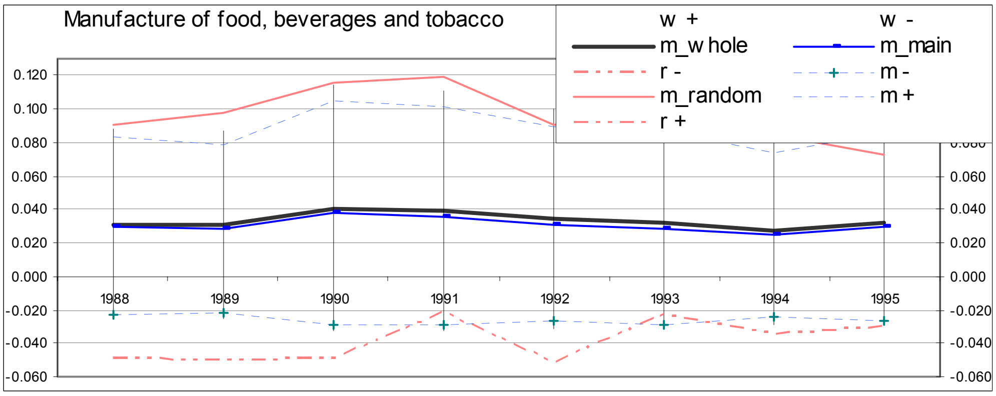

Manufacture of metals
w +
m_whole
r -
m_random
r +
w -
m_main
m -
m +
0.140
0.120
0.100
0.080
0.060
0.040
0.020
0.000
-0.020
-0.040
-0.060
1988
1989
1990
1991
1992
1993
1994
1995

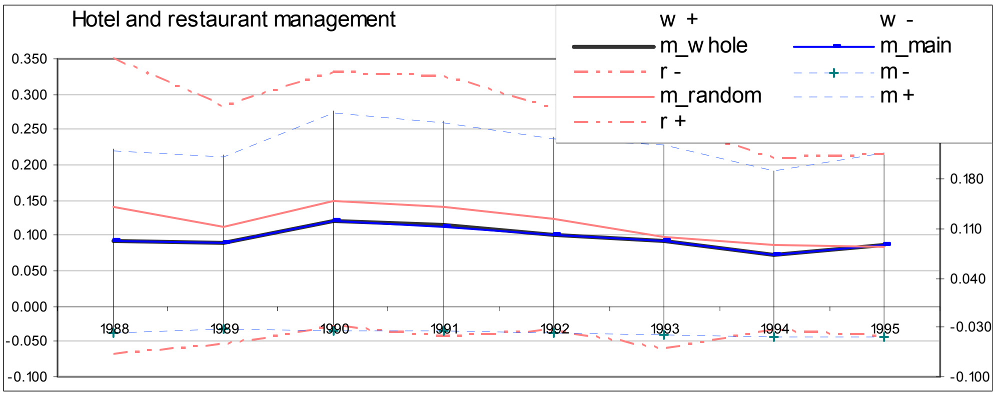

Manufacture of oil rigs
w +
m_whole
r -
m_random
r +
w -
m_main
m -
m +
0.180
0.130
0.080
0.030
-0.020
-0.070
1988
1989
1990
1991
1992
1993
1994
1995


Other industry production
w +
w -
m_w hole
m_main
r -
m -
m_random
m +
r +
0.150
0.100
0.050
0.000
-0.050
1988
1989
1990
1991
1992
1993
1994
1995

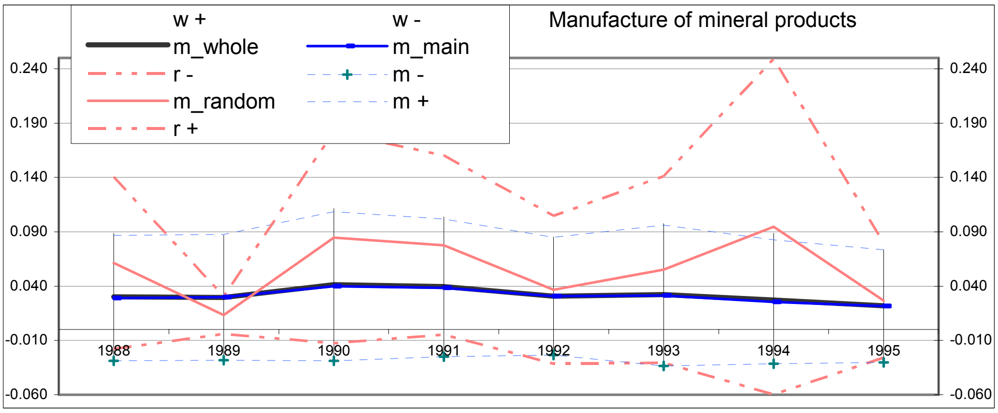

Other engineering production
w +
w -
m_whole
m_main
r -
m -
m_random
m +
r +
0.120
0.100
0.080
0.060
0.040
0.020
0.000
-0.020
-0.040
1988
1989
1990
1991
1992
1993
1994
1995

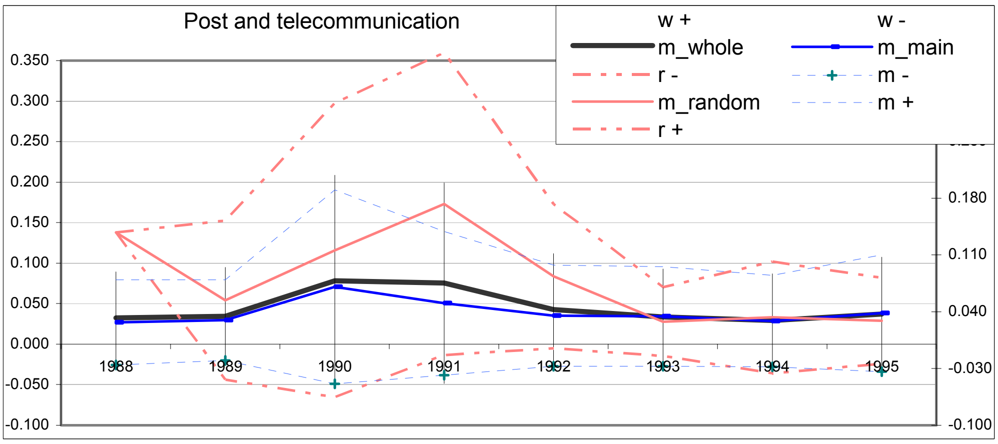

Construction of crafts and boats
w +
w -
m_w hole
m_main
r -
m -
m_random
m +
r +
0.120
0.100
0.080
0.060
0.040
0.020
0.000
-0.020
-0.040
1988
1989
1990
1991
1992
1993
1994
1995


Manufacture of food, beverages and tobacco
w +
w -
m_w hole
m_main
r -
m -
m_random
m +
r +
0.120
0.100
0.080
0.060
0.040
0.020
0.000
-0.020
-0.040
-0.060
1988
1989
1990
1991
1992
1993
1994
1995



The same pattern is valid for the industry Transport and storage (code 911).





For the industries Wood conversion (code 321), Manufacture of chemicals (code 331), Manufacture of metals (code 341) and Post and telecommunication (code 640) the random part constitutes correspondently 25%, 27%, 60% and 13% of total observations. For these industries, average *p* and its standard deviation increases after the inclusion of the random part with higher bankruptcy probabilities. However, in general influence on *p* is not large. For industries Manufacture of oil rigs (code 461) and Other industry production (code 490), with correspondently 65.5% and 49% of the random part with a lower *p*, we observe a small decrease in the average *p* and its standard deviation. For the rest of the industries considered: industry Construction of crafts and boats (code 462), Other engineering production (code 469), Transport and storage (code 911) with correspondent random part of 54%, 47% and 19% and some other industries, there were almost no change in the values of *p*. This is true also for the industry Manufacture of food, beverages and tobacco (code 411), which observations are mixed with observations on risky enterprises from industry Hotel and restaurant management (code 550). However, for these industries, the number of randomly added observations is quite small (4,66 %), so influence is not large. Consequently, inclusion of the arbitrary part does not have any considerable effect on the bankruptcy probabilities. The same pattern is valid for the industry Transport and storage (code 911).

Combining of information on bank loans and risk indicators:

After establishing the correspondence between industry codes in the bank statistics and industry codes for the individual enterprises, we proceed with aggregation of the individual risk indicators and weighting them with loan volumes. Referring to the two alternative aggregation dimensions, we use industry/year groups, i.e. the aggregate across all counties, and industry/county/year groups. The first type of aggregation mixes observations across counties and can be in disagreement with the county specific type of activities of the medium-size savings banks. However, it provides a direct link between the two datasets. Moreover, it may be more accurate than the second one if banks in their annual reports assign counties on some other basis (e.g. location of the local branch which the enterprise use for the loan application), than the formal registration criteria used in the SEBRA-database. The second type of aggregation allows utilization of higher variation in risk indicators (i.e. over larger number of groups). However, it will underestimate risks for some banks if we assume zero risk exposure for those industry/county/year groups that have missing risk measures in all directions of bank's lending. Therefore, we need some adjustments for the groups with missing aggregate bankruptcy probabilities and the following approaches are suggested:

1. Narrow the sample and calculate bankruptcy probabilities for a smaller number of industry/county/year groups, i.e. joining or excluding some counties;
2. Interpolating accounting data, and thus obtain estimated and predicted bankruptcy probabilities without gaps;
3. Interpolating bankruptcy probabilities to cover the gaps for some years;
4. Map bankruptcy probabilities for the missing groups from the similar ones, e.g. groups for the same industry and neighbouring counties and possibly central Oslo/Akershus county)

Bias in the final risk indicators for the banks arising due to the suggested adjustments is not expected to be large, because in most cases shares of loans with missing risk measures constitute only less than 1 per cent banks' total loans. Only Soknedal Sparebank (bank number 4333) has a particular high share of loans which corresponds to the industry/county/year groups with missing aggregate bankruptcy probabilities. It has loans to different industries, including Collection, purification and distribution of water (code 410), but only in county 16 (Sør-Trøndelag) according to the banks statistics. While industry 410 has only few enterprises registered in this county and none of them submitted accounting information for the years 1996 and 1998-2000. Thus, aggregate bankruptcy probabilities are missing for the groups 410/16/1996, 410/16/1998 and 410/16/2000. In total, share of loans for which it is not possible to obtain risk measure reaches almost 12 per cent for this bank in 2000. Shares of 1 to 3 per cent are attained for the following banks. Totens Sparebank (bank number 2050) provides loans almost only to counties 3 and 5 according to the bank statistics, while none of the enterprises in industry Inland water transport (code 612) are registered in county 5. Kragerø Sparebank (bank number 2655) grants loans almost only to county 8; while none of the enterprises in industry Fishing (code 051) are registered in this county. Nordmøre Sparebank (bank number 3930) has loans almost only to county 15 (Møre and Romsdal), except of Oslo/Akershus, and some of the enterprises in industry 410 are registered there but none of them submitted accounting information for the year 1999. Sparebank Midt-Norge (bank number 4201) provides loans mainly to counties 16 and 17. However, the SEBRA-database does not contain accounting information on enterprises registered in industry Extraction of crude, petroleum and natural gas (code 110) and county 17, in industries 410 and 950 and county 16. Also enterprises in industry 051 and county 17 did not submit accounting information for the year 1996. A similar problem appears for the Melhus Sparebank (bank number 4230), Meldal Sparebank (bank number 4260) and Sparebank Hemne (bank number 4312), which provide loans mainly to the enterprises in county 16 according to the bank statistics, therefore it is not possible to obtain risk measures for industry Collection, purification and distribution of water (code 410), Mining of coal and lignite (code 100) and Tanning and dressing of leather (code190) for the years 1998 - 2000.

These examples suggest that procedure of joining and excluding some counties is not very efficient because missing aggregate bankruptcy probabilities correspond to a wide array of counties and industries and usually due to the irregularities in the accounting information. At the same time, interpolation of the accounting data will provide biased results in the estimation of the bankruptcy probabilities and will not solve the problem due to the absence of any information for some of the industry/county groups. Therefore, we apply mapping for the missing aggregate bankruptcy probabilities from similar groups or from the aggregate industry vector across all counties. Below are some examples of missing aggregate bankruptcy probabilities for the particular industries and counties:

- For industry 100 and counties 9, 14, 15, 16, 18, p is not available. Since data is available only for the counties 1, 3, 17, 18 and 21, average p (excluding county 21) can be taken as a proxy.
- For industry 140 and county 21, p is not available. We can use p from neighbouring county 20.
- For industry 160 and county 10: p is available only for county 3 and can be taken as proxy.
- For industry 221, data is available only for county 11 and a correspondent p is taken.
- For industry 410 and counties 8, 14, 16, 19, 11, 18, 20, p is not available. So we can use average p for the available counties 3, 4, 5, 6, 7, 10, 11, 12, 17, 18.
- For industry 540 and county 17, we can use average p for the neighbouring counties 16, 18 and 19.

Substitution of average p is calculated for the following neighbouring counties[39]:

```
For county 1, use counties 3,4,5,6,7;
For county 3, use counties 1,4,5,6,7;
For county 4, use counties 1,3,5,6,7,16;
For county 5, use counties 1,3,4,5,6,7,8,9,10,11,12,14,15,16;
For county 6, use counties 1,3,5,7,8,9;
For county 7, use counties 1,3,6,8,9;
For county 8, use counties 6,7,9,10,11;
For county 9, use counties 7,8,10,11;
For county 10, use counties 8,9,11,12;
For county 11, use counties 9,10,12;
For county 12, use counties 6,8,11,14,15;
```

[39] List of counties in Norway: 1 Østfold, 2 Akershus, 3 Oslo, 4 Hendmark, 5 Oppland, 6 Buskerud, 7 Vestfold, 8 Telemark, 9 Aust-Agder, 10 Vest-Agdler, 11 Rogaland, 12 Hordaland, 14 Sogn and Fjordane, 15 Møre and Romsdal, 16 Sør-Trondelag, 17 Nord-Trondelag, 18 Nordland, 19 Troms, 20 Finnmark, 21 Svalbard

```
For county 14, use counties 5,6,12,15,16;
For county 15, use counties 5,12,14,16;
For county 16, use counties 4,5,15,17;
For county 17, use counties 16,18,19;
For county 18, use counties 17,19,20;
For county 19, use counties 18,20,17;
For county 20, use counties 18,19,17;
For county 21, use counties 19 and 20.
```

When less than two values of p are available for a particular industry in the specified neighbouring region, average *p* for all counties is used (e.g. industries: 100, 160, 190, 221, 410, 603, 640 and 950).

By this procedure we obtain proxies for bankruptcy probabilities for the industry/county/year groups that are constructed using statistics on loans but are not directly obtainable from the SEBRA-database. Then a risk indicator for the bank is calculated as a weighted average of all relevant for its loan portfolios bankruptcy probabilities. For each of the years 1998-2001, there are maximum 30 banks (around 1/5 of all banks in the database) with risk indicator based on proxies for missing bankruptcy probabilities. Most of them are either large banks, with highly diversified loan portfolio and loans for different industries and counties (for example banks 1802, 3201, 4701, 6001, 7001 and 8902), or middle size savings banks, which have loans in only one or two counties (for example banks 3450 and 4201). The share of these banks in the total volume of loans varies from 14 to 18 per cent over the years and it does not exceed 0.1 per cent in the volume of loans to the groups with missing aggregate bankruptcy probabilities. Therefore, contracted proxies for bankruptcy probabilities do not have large influence on the resulted risk indicators for the banks and this fact favours using industry/county/year groups.

WORKING PAPERS (ANO) FROM NORGES BANK 2002-2004

Working Papers were previously issued as Arbeidsnotater from Norges Bank, see Norges Bank's website http://www.norges-bank.no

2002/1 Bache, Ida Wolden
Empirical Modelling of Norwegian Import Prices
Research Department 2002, 44p

2002/2 Bårdsen, Gunnar og Ragnar Nymoen
Rente og inflasjon
Forskningsavdelingen 2002, 24s

2002/3 Rakkestad, Ketil Johan
Estimering av indikatorer for volatilitet
Avdeling for Verdipapirer og internasjonal finans Norges Bank 33s

2002/4 Akram, Qaisar Farooq
PPP in the medium run despite oil shocks: The case of Norway
Research Department 2002, 34p

2002/5 Bårdsen, Gunnar, Eilev S. Jansen og Ragnar Nymoen
Testing the New Keynesian Phillips curve
Research Department 2002, 38p

2002/6 Lindquist, Kjersti-Gro
The Effect of New Technology in Payment Services on Banks'Intermediation
Research Department 2002, 28p

2002/7 Sparrman, Victoria
Kan pengepolitikken påvirke koordineringsgraden i lønnsdannelsen? En empirisk analyse.
Forskningsavdelingen 2002, 44s

2002/8 Holden, Steinar
The costs of price stability - downward nominal wage rigidity in Europe
Research Department 2002, 43p

2002/9 Leitemo, Kai and Ingunn Lønning
Simple Monetary Policymaking without the Output Gap
Research Department 2002, 29p

2002/10 Leitemo, Kai
Inflation Targeting Rules: History-Dependent or Forward-Looking?
Research Department 2002, 12p

2002/11 Claussen, Carl Andreas
Persistent inefficient redistribution
International Department 2002, 19p

2002/12 Næs, Randi and Johannes A. Skjeltorp
Equity Trading by Institutional Investors: Evidence on Order Submission Strategies
Research Department 2002, 51p

2002/13 Syrdal, Stig Arild
A Study of Implied Risk-Neutral Density Functions in the Norwegian Option Market
Securities Markets and International Finance Department 2002, 104p

2002/14 Holden, Steinar and John C. Driscoll
A Note on Inflation Persistence
Research Department 2002, 12p

2002/15 Driscoll, John C. and Steinar Holden
Coordination, Fair Treatment and Inflation Persistence
Research Department 2002, 40p

2003/1 Erlandsen, Solveig
Age structure effects and consumption in Norway, 1968(3) – 1998(4)
Research Department 2003, 27p

2003/2 Bakke, Bjørn og Asbjørn Enge
Risiko i det norske betalingssystemet
Avdeling for finansiell infrastruktur og betalingssystemer 2003, 15s

2003/3 Matsen, Egil and Ragnar Torvik
Optimal Dutch Disease
Research Department 2003, 26p

2003/4 Bache, Ida Wolden
Critical Realism and Econometrics
Research Department 2002, 18p

2003/5 David B. Humphrey and Bent Vale
Scale economies, bank mergers, and electronic payments: A spline function approach
Research Department 2003, 34p

2003/6 Harald Moen
Nåverdien av statens investeringer i og støtte til norske banker
Avdeling for finansiell analyse og struktur 2003, 24s

2003/7 Geir H. Bjønnes, Dagfinn Rime and Haakon O.Aa. Solheim
Volume and volatility in the FX market: Does it matter who you are?
Research Department 2003, 24p

2003/8 Olaf Gresvik and Grete Øwre
Costs and Income in the Norwegian Payment System 2001. An application of the Activity Based Costing framework
Financial Infrastructure and Payment Systems Department 2003, 51p

2003/9 Randi Næs and Johannes A.Skjeltorp
Volume Strategic Investor Behaviour and the Volume-Volatility Relation in Equity Markets
Research Department 2003, 43p

2003/10 Geir Høidal Bjønnes and Dagfinn Rime
Dealer Behavior and Trading Systems in Foreign Exchange Markets
Research Department 2003, 32p

2003/11 Kjersti-Gro Lindquist
Banks' buffer capital: How important is risk
Research Department 2003, 31p

2004/1 Tommy Sveen and Lutz Weinke
Pitfalls in the Modelling of Forward-Looking Price Setting and Investment Decisions
Research Department 2004, 27p

2004/2 Olga Andreeva
Aggregate bankruptcy probabilities and their role in explaining banks' loan losses
Research Department 2004, 44p

KEYWORDS:

Bank losses
Bankruptcy probabilities
Aggregation

Lobo Medio - 17627