

# Corporate Credit Risk Modeling and the Macroeconomy<sup>‡</sup>

Kenneth Carling Tor Jacobson Jesper Lindé Kasper Roszbach<sup>‡</sup>

5 April 2006

## Abstract

Despite a surge in the research efforts put into modelling credit risk during the past decade, few studies have incorporated the impact that macroeconomic conditions have on business defaults. In this paper, we estimate a duration model to explain the survival time to default for borrowers in the business loan portfolio of a major Swedish bank over the period 1994-2000. The model takes both firm-specific characteristics, such as accounting ratios and payment behaviour, loan-related information, and the prevailing macroeconomic conditions into account. The output gap, the yield curve and consumers' expectations of future economic development have significant explanatory power for the default risk of firms. We also compare our model with a frequently used model of firm default risk that conditions only on firm-specific information. The comparison shows that while the latter model can make a reasonably accurate ranking of firms' according to default risk, our model, by taking macro conditions into account, is also able to account for the absolute level of (default, and thus also credit) risk.

**Keywords:** Firm default, default risk, credit risk, corporate loans, duration model, survival, macroeconomy.

**JEL:** C23, C24, C25, C41, E44, G21, G33, L25.

---

\*We are grateful to two anonymous referees and the editor for their helpful comments. Thanks to Daniel Hammarsten and Daniel Sunesson for excellent research assistance. We are also indebted to Sonja Daltung, Frank Heid, Gunnar Isaksson, Olivier Renault, Anders Vredin and seminar participants at the Federal Reserve Bank of Philadelphia, the BIS research task force workshop on Applied Banking, the Society for Economic Dynamics 2001 meeting, the LSE-FMG European Investment Review conference 2002, the SIFR conference on The Theory and Practice of Credit Risk Modelling and the OCC conference on Credit Rating and Scoring Models for comments and suggestions. All remaining shortcomings are the authors' responsibility. An early version of this paper circulated under the title "Capital Charges under Basel II: Corporate Credit Risk Modelling and the Macro Economy".

<sup>†</sup>The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank.

<sup>‡</sup>Carling is at IFAU and Dalarna University. *Email:* kca@du.se. The other authors are at the Research Division, Sveriges Riksbank, SE 103 37 Stockholm. *Email:* firstname.lastname@riksbank.se.

# 1 Introduction

The literature on default and credit risk modeling is extensive and growing. A pioneering contribution from the 1960's is Altman's study of business default risk [4]. Following Altman, a number of authors have estimated various types of default risk models on cross-sectional data sets. For example Altman [5] [6], Frydman, Altman and Kao [17], Li [29], and Shumway [37], who all focus on the analysis of bankruptcy risk at the firm level.

In the last decade, a whole range of modeling techniques has been developed to analyze *portfolio* risk. Broadly viewed, there are four groups of portfolio credit risk models. The first group is 'structural' and based on Merton's [36] model of firm capital structure: individual firms default when their assets' value fall below the value of their liabilities. The second group consists of econometric factor risk models, where default risk in 'homogeneous' subgroups is determined by a macroeconomic index and a number of idiosyncratic factors. Both groups consist of 'bottom-up' models that compute default rates at either the individual firm level or at sub-portfolio level. The third group contains 'top-down' actuarial models, like Credit Suisse's CreditRisk+, that make no assumptions regarding causality. Finally, a number of authors, such as Carey [10], use non-parametric methods.

Koyluoglu and Hickman [26] provide a description of the above mentioned types of portfolio credit risk models. They note that all model types, despite their differences, are built on three more or less general components to calculate portfolio loss distributions. First, they contain some process that generates conditional default rates for each borrower in each state of nature and a measure of covariation between borrowers in different states of nature. Second, their set-up allows for the calculation of conditional default rate distributions for sets of homogeneous sub-portfolios (e.g., rating classes) as if individual borrower defaults are independent, since all joint behavior is accounted for in generating conditional default rates. Third, unconditional portfolio default distributions are obtained by aggregating homogeneous sub-portfolios' conditional distributions in each state of nature; then conditional distributions are averaged using the probability of a state of nature as the weighting factor. Gordy [21] confirms the insights of Koyluoglu and Hickman in a comparison of two influential benchmark credit risk models, CreditMetrics and CreditRisk+. He concludes that they have similar mathematical structures and that the main sources of discrepancies in predictions are differences in distributional assumptions and functional forms.

Credit risk models are also used to calculate capital requirements for banks. Estrella [13], for example, constructs a theoretical model of optimal bank capital. He finds that a regulatory minimum capital requirement based on VaR is likely to be procyclical and suggests some ways to remedy this procyclicality. Gordy [22] examines the relation between portfolio models of credit VaR and ratings-based bucket models. He concludes that the latter can be reconciled with the

general class of credit VaR models. Calem and LaCour-Little [9] estimate a survival time model for mortgage loan data and use Carey’s [10] method to simulate the distributions of default probabilities (PD’s). They find that capital charges are generally below the current standard - thereby providing some empirical support for the occurrence of securitization. Hamerle, Liebich and Rösch [23] follow another approach and model the (unconditional) PD’s by means of a non-linear random effects probit and logit model.

A number of studies investigate the relevance of macroeconomic conditions for the estimation of credit risk. Allen and Saunders [2] survey the literature on how to incorporate systemic influences into risk measurement. Wilson [38], outlines the general principles behind McKinsey’s proprietary portfolio credit risk model, CreditPortfolioView, incorporating a set of macroeconomic variables in a multifactor country/sector logit model. No information is published, however, on the coefficient estimates and other model properties. Nickell, Perraudin and Varotto [32] use macro dummies in an ordered probit model of corporate bond rating transitions. Other approaches use the transition matrices to generate a linkage between underlying macroeconomic conditions and asset quality (Bangia, Diebold, Kronimus, Schagen and Schuermann [8]) or Monte Carlo resampling methods to generate unconditional (in terms of macroeconomic conditions) portfolio credit loss distributions (e.g. Carey [10]). Pesaran, Schuermann, Treutler and Weiner [33] show how a global macroeconometric model can be linked to firm-specific returns in a Merton-type credit risk model. In their paper, firm credit risk is made contingent on variables such as changes in equity indices, interest rates, inflation, real money balances, oil prices and output. Pesaran, Schuermann and Treutler [34] extend this model and study the potential for credit risk diversification.

This paper contributes to the field by presenting a duration model for the survival time until default of business-loan borrowers that includes both firm-specific and macroeconomic explanatory variables. Such a model can be used not only for analyzing business defaults but also as an input in a portfolio credit risk model. We employ a rich data set from a leading Swedish retail bank that is also internationally active. The reduced form model enables us to enter factors that drive default behavior in the corporate sector and to quantify to what extent they contribute to predicting default realizations. Our approach is similar to the one used in the above mentioned group of econometric factor risk models, but differs in some of its assumptions. We allow, for example, explicitly for duration dependence. Also, by including macro explanatory variables, we increase the number of sources of inter-firm default risk correlation. An important reason for using a reduced form model is that, contrary to previous studies of corporate default risk, we have access to of a large longitudinal data set with observations on more than 50,000 firms during 24 quarters. This bank data has been augmented with real time information on the characteristics of the firms from a leading credit bureau in Sweden, Upplysningscentralen AB (hereafter: UC). As a result of this, we can separate idiosyncratic and macroeconomic effects by

exploring the time-series and the cross-sectional dimensions of the data. Since an overwhelming majority of these firms are not publicly quoted, applying a stock-price based model is not feasible. Estimating a micro model of default risk using financial statement data is therefore a logical approach.

Our main finding is that macroeconomic variables have significant explanatory power for firm default risk in addition to a number of common financial ratios. Both the output gap, the yield curve and households' expectations about the Swedish economy are quantitatively important indicators of the evolution of default risk over time. Their quantitative importance is also reflected in a substantially improved fit. Our model, by taking macro conditions into account, manages to capture the absolute level of default risk, while benchmark models, conditioning only on firm-specific information, can only make reasonably accurate *rankings* of firms according to their default risks.

We also find evidence for the existence of duration dependence. This implies that a binary model specification is inappropriate and that idiosyncratic risk factors need to be complemented with information on survival time to obtain consistent default-risk estimates. Finally, we find that the risk of default is markedly higher for short-term loans than for long-term credit. The risk of default is also markedly higher for short-term loans than for long-term credit. This is in accordance with a common thought that banks are more inclined to extend short-term loans to firms with a more uncertain future, while safer enterprises tend to have better access to loans with longer, fixed, maturities.

The remainder of the paper is organized as follows. Section 2 offers a detailed description of the data. In Section 3 we present the statistical model and the empirical estimates. Section 4 concludes.

## 2 Data

The data set we use is a panel consisting of 579,941 observations on the bank's business customers. It amounts to six years of quarterly data on all 54,603 Swedish *aktiebolag* companies that had one or several loans outstanding at the bank on the last day of at least one quarter between March 31, 1994, and March 31, 2000. *Aktiebolag* are by approximation the Swedish equivalent of US corporations and UK limited businesses. Swedish law requires every *aktiebolag* to have at least SEK 100,000 (approximately US\$ 14,000) of equity to be eligible for registration at the Swedish Patent and Registration Office (PRV). Firms are also required to submit an annual report to PRV. Although we have annual report data on small firms such as general partnerships, limited partnerships and sole proprietors, these will be disregarded because we do not have access to the relevant credit histories. This means that we have deleted approximately

20% of all firms in the portfolio.<sup>1</sup> Observe, however, that a large part of the sample still consists of relatively small enterprises: 65% of the observations concern businesses with 5 or fewer employees.

The data on these firms come from two sources: from the bank itself and from the credit bureau, UC. The bank supplied a full history of internal credit related data, including variables like the amount of credit granted, actual exposure, the types of credit, the amount granted per credit type, collateral, payment status, and an internal risk classification. These data are available at a quarterly frequency. Upplysningscentralen provided us with non-bank specific data for each company in the bank's portfolio, which it collects from the PRV annual report data. For example, balance sheet and income statement data from the annual report were provided, but also historical data on events related to firms' payment remarks. These data were available at different frequencies, varying from daily for payment remarks to annually for accounting data. We will discuss both data sources in greater detail below.

## 2.1 Bank data

As part of its risk management system the bank maintains an internal rating scheme; this requires each business customer to be assigned to one of 15 credit rating grades. Rating class 1 represents the highest credit quality and rating grade 15 stands for the lowest credit quality, factual default, with the intermediate rating classes intended to imply a monotonically increasing risk profile. At a minimum, the bank updates the credit rating of each firm in its portfolio every 12 months. For the purpose of this study we will use the bank's definition of a default: a borrower that is assigned to rating class 15. The necessary condition to be assigned to rating class 15 is that (i) principal or interest payments are 60 days overdue, and (ii) a bank official has to make a judgement and reach the conclusion that any such payment is unlikely to occur in the future. A comparison with data from the credit bureau shows that rating class 15 is highly correlated with (the officially registered legal concept of) bankruptcy. Generally, default by rating class 15 leads the latter by one or more quarters, most likely due to the complexity of legal procedures that have to be completed before bankruptcy is officially invoked.

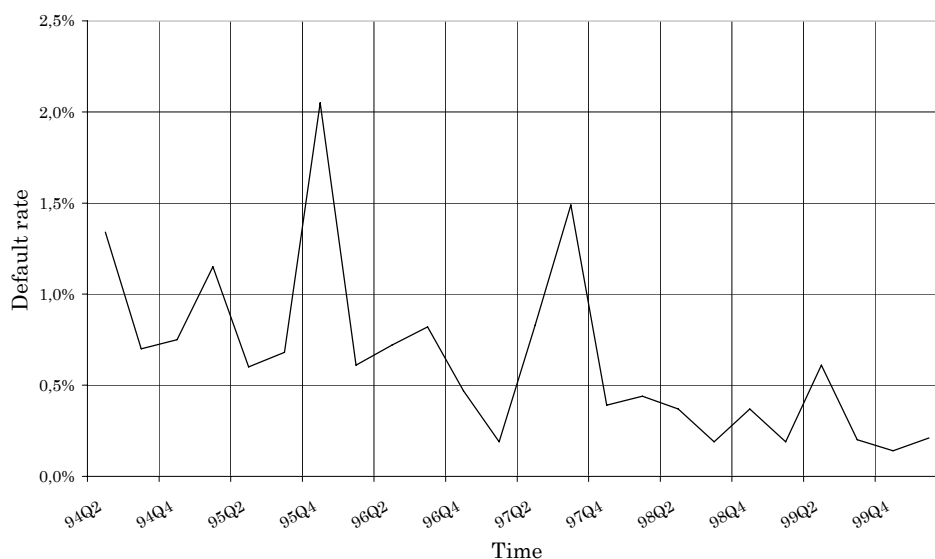
Figure 1 shows that there is quite some movement over time in the average default rate of the bank's portfolio. The maximum rate of default within the sample period is reached in the fourth quarter of 1995, at a level of 2%. Over the whole of 1995, 4% of the bank's borrowers defaulted, compared with an average annual rate over the whole sample period of 2.5%. After 1995 the default rate declines, reaches a smaller peak in the third quarter of 1997 of 1.5%, and then steadily falls to very low levels towards the end of the sample period.

The bank provided us with the complete credit history of each business customer. The more important bank variables are: the size of the approved credit lines, actual exposure, the rating

---

<sup>1</sup>However, these 20% of firms only represent approximately 6% of the portfolio loan value.

Figure 1: Quarterly default rates for the whole portfolio.



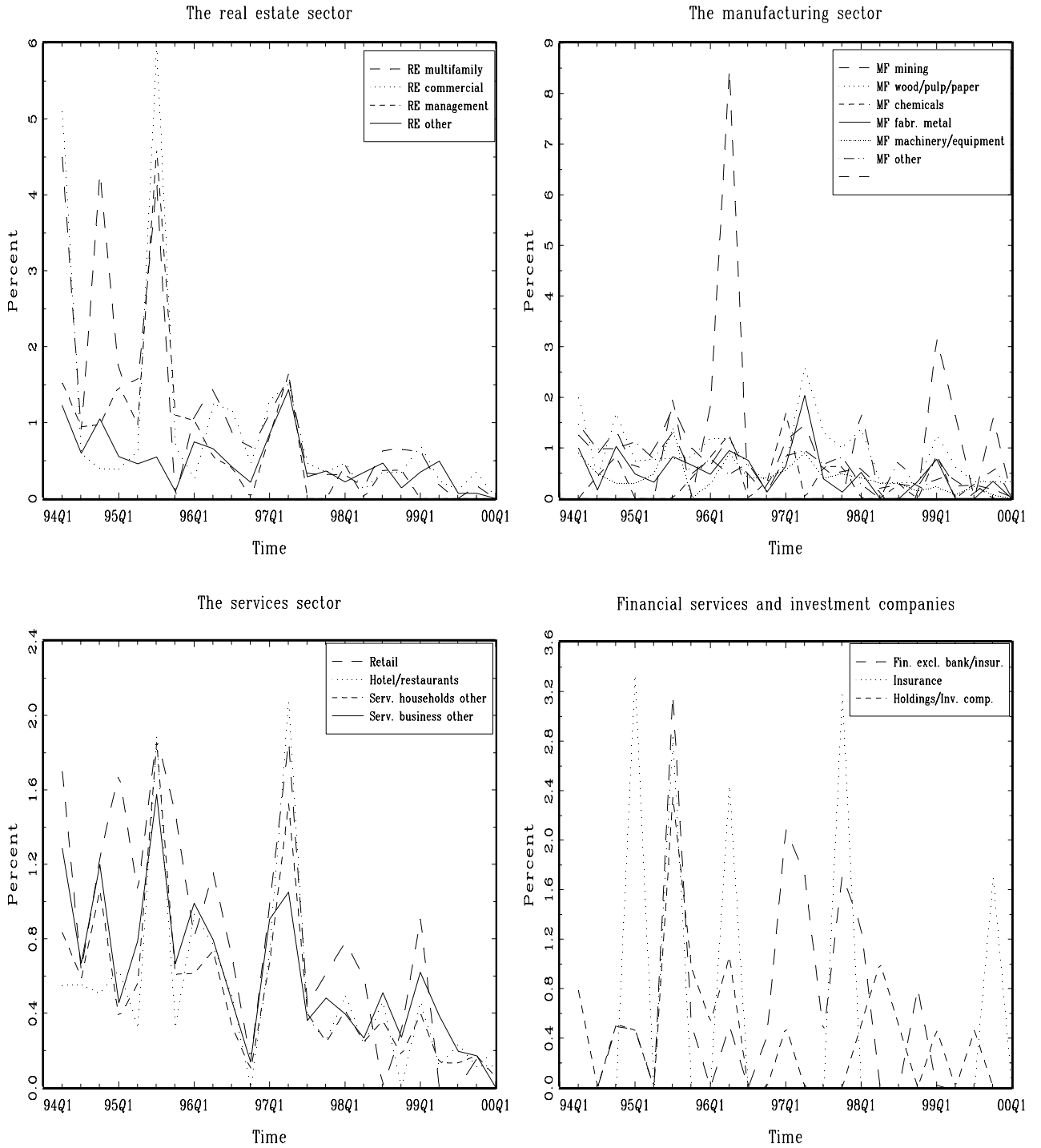
class, the industry code, and a number of variables splitting up total credit in different types of loans. Appendix A contains a full list of the variables provided to us by the bank. We reduced a total of 19 types of credit to 5 broader groups, also used by the bank for certain analytical purposes: short term lending, long term lending, mortgages, guarantee loans and the remainder, mixed loans. Of all customer observations, 67% involve short term loans while 32% concern long run loans, 5% mortgages, 17% guarantee loans and 20% mixed loans (the remaining credit types). More than 40% of the observations involve at least two types of credit, and in 21%, businesses have both a short and a long term loan, implying that about two thirds of the businesses that borrow long also borrow short term. Other credit type combinations that have a frequency of 5% or more are: short term and guarantee loans, short term and mixed loans and guarantee and mixed loans. The average bank relationship duration for a firm is 10.8 quarters.<sup>2</sup> If split up according to credit type, the average duration for a short term loans is 9.9 quarters, whereas long term loans have an average duration of 10.7 quarters.

Figure 2 shows that the default rate varies significantly, not only over time, but also between, and even within, industries. For most of the sample period, short term loans are associated with the highest default rates. One exception is the last quarter of 1994, when mortgage defaults reach a peak of 3.1% and the period 1997 Q3 - 1998 Q3, when the long-term loan default rates slightly exceed short-term rates. The four largest industries in terms of total average exposure are multi-family real estate, manufacturing of machinery & equipment, commercial real estate and wholesale. Together they account for 48% of the bank's business loan portfolio

---

<sup>2</sup>This includes censored observations.

Figure 2: Default rates by industry.



on average. The quarterly default rates in three of these industries peak simultaneously in the third quarter of 1995, although at highly varying levels. The fourth industry, multi-family real estate, reaches its peak default rate in the same quarter as the mortgage default rate: 1994 Q4. Commercial and multi-family real estate have the highest quarterly shares of defaults, 6.0% and 4.1% respectively. Wholesale and machinery & equipment default rates only reach top levels of 1.6% and 1.4%. After 1995, all four industries more or less follow the economy-wide pattern, their peaks in 1996 Q2 and 1997 Q2 ranging from .9% to 1.6%. Most other industries display a similar pattern over time. Two exceptions are the services industry, where the default rate appears to be more persistent, and the financial services industry, which displays a more erratic behavior, probably because of the smaller number of borrowers. In terms of average default rates, the commercial and multi-family real estate sectors rank first and second with quarterly rates of over 1%, followed at short distance by mining & quarrying, wood, pulp & paper and hotel & restaurants with rates between .8 and .9%. The best performing industries are electricity/gas and banking, with average rates of .2 and 0%. Chemicals, machinery & equipment and transport are the only other sectors with average default rates below 0.5%.

## 2.2 Credit bureau data

The data from the credit bureau contains information on most standard balance sheet and income statement variables. Some examples of balance sheet entries are cash, accounts receivable and payable, current assets and liabilities, fixed and total assets, total liabilities and total equity. Some examples of the income statement entries are total turnover, earnings before interest, depreciation and amortization, depreciation, financial income, extraordinary income and taxes. Appendix B contains a complete list of all annual report variables. In addition to the annual report data collected by the Swedish Patent and Registration Office, we have information on the firms' track records regarding payment behavior as recorded by remarks for 61 different credit and tax related events.<sup>3</sup> Two types of remarks exist. The first type is non-payment remarks, the storage and usage of which are regulated by the Credit Information Act, the Personal Data Act and overseen by the Swedish Data Inspection Board. Examples of events that are registered are: delays in tax payments, the repossession of delivered goods, the seizure of property, the resettlement of loans and actual bankruptcy. In practice, with a record of non-payment remarks individuals will not be granted any new loans and businesses will find it more difficult to open new lines of credit. The second type is bank specific remarks, reflecting a firm's payment behavior with banks. All Swedish banks participate in this scheme and report any abuse of a bank account or a credit card and slow loans (loans of which repayment is considered questionable) to the credit bureau that maintains these records. Their storage and usage is only

---

<sup>3</sup>In 2004 the Patent Office was split up into two separate organizations; annual report data are now collected by Bolagsverket (the Swedish Companies Registration Office).

regulated by the Personal Data Act. Whereas a bank remark may have the same consequences as a non-payment remark, this is not generally the case. Their effect on individual applications for credit presumably works mainly through the accumulation of negative indicators.

**Table 1: Descriptive statistics for the credit bureau data.**

Statistics are for truncated variables [at 1% and 99% percentiles]

Spell	Unit	$N$	$\mu$	$\sigma$	min	1%	50%	99%	max
<b>Performing</b>		576343							
TS	(mn SEK)		19.66	49.39	0	.02	2.94	250.00	250.00
EBITDA/TA	(ratio)		0.11	.19	-.80	-.61	.11	.66	.80
TL / TA	(ratio)		.71	.23	0	.03	.76	1.00	1.00
I / TS	(ratio)		.10	.19	0	.00	.02	1.00	1.00
Bank pay-remark	(share)		<.01	.05	0				1
Legal pay-remark	(share)		.01	.10	0				1
<b>Defaulted</b>		3598							
TS	(mn SEK)		8.49	25.44	0	.00	1.91	124.52	250.00
EBITDA/TA	(ratio)		.01	.25	-.80	-.80	.04	.70	.80
TL / TA	(ratio)		.84	.23	0	.00	.93	1.00	1.00
I / TS	(ratio)		.14	.23	0	.00	.04	1.00	1.00
Bank pay-remark	(share)		.17	.38	0				1
Legal pay-remark	(share)		.30	.46	0				1

Notes: The definition of variables are: TS = total sales; EBITDA = earnings before interest payments, taxes, depreciation and amortizations; TA = total assets; TL = total liabilities; I = inventories; Bank pay-remark = a dummy variable taking the value of 1 if the firm has a “non-performing loan” at a bank in the preceding four quarters; Legal pay-remark = a dummy variable taking the value of 1 if the firm has a payment remark due to one or more of the following events in the preceding four quarters; a bankruptcy petition, issuance of a court order to pay a debt, seizure of property. All variables are in nominal terms and in million SEK.

As can be seen in Table 1, all descriptive statistics for accounting ratios and other credit bureau variables, such as non-payment and bank remarks and sales, were calculated based on different numbers of observations. For various reasons and depending on the specific variable, a maximum of 28,000 observations per variable, approximately 5 percent of the data, could not be used directly in the estimation of the model. This could be due to incorrect entering of data by the credit bureau (unreasonable or negative values for non-negative balance sheet and income statement variables like total liabilities, total assets, inventories and sales), because of the nature of the ratio (a zero in the denominator), or simply the absence of an entry (some companies fail to submit their annual report). In all, this would have implied the deletion of approximately 10% of the sample. To avoid such reduction of our sample size, we replaced missing data entries

by the firm-specific mean of that variable over time. As a result, the final estimation is done with the full sample of 54,603 firms observed on 579,941 occasions.<sup>4</sup>

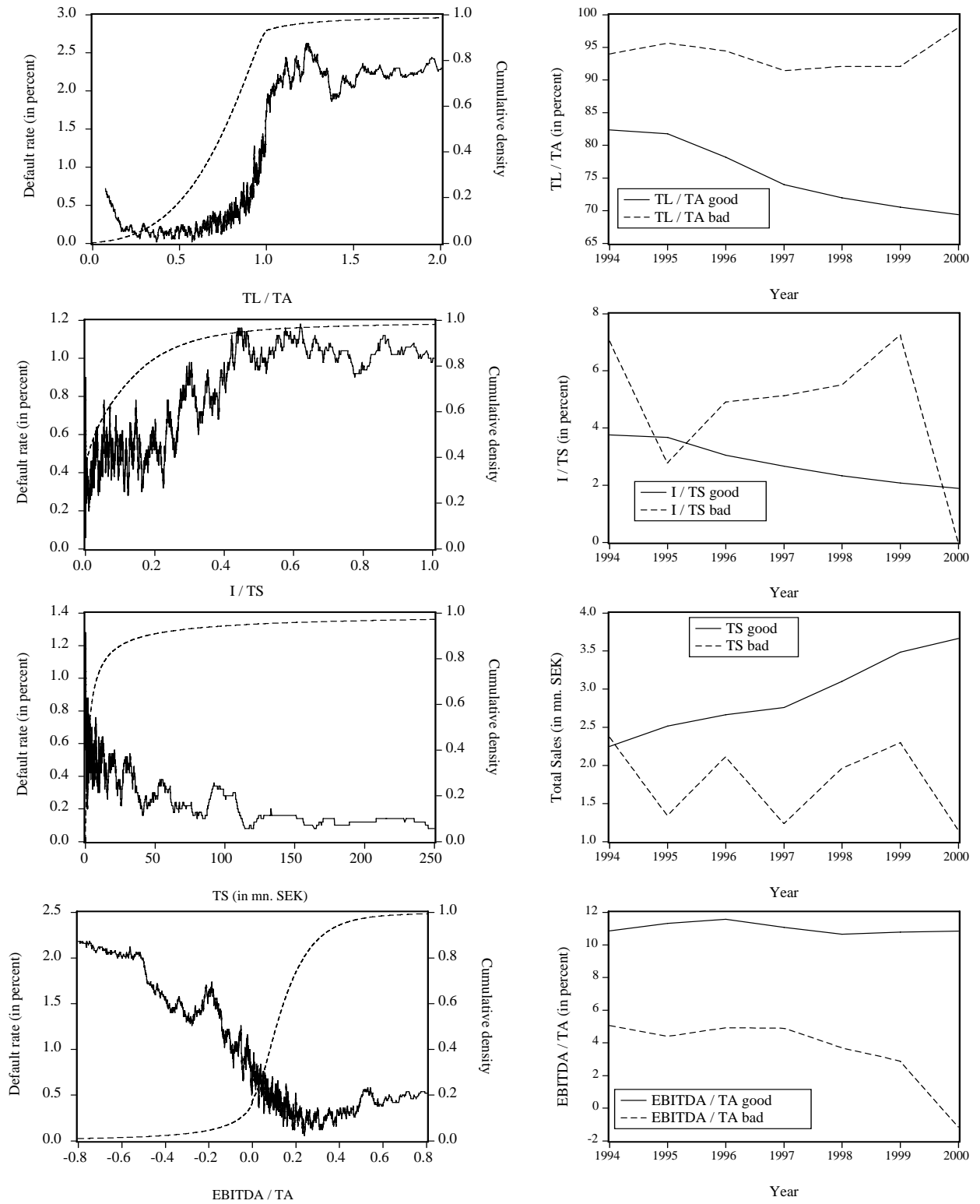
As annual reports typically become available with a significant time lag, it cannot in general be assumed that accounting data for year  $\tau$  were available during or even at the end of year  $\tau$  to forecast default risk in year  $\tau + 1$ . To account for this, we have lagged all accounting data by 4 quarters in the estimations. For most companies, who report balance sheet and income data over calendar years, this means that data for year  $\tau$  is assumed to have been available in the first quarter of year  $\tau + 2$ . For a number of firms some transformation had to be applied to the accounting variables to adjust for reporting periods that did not coincide with the calendar year, to assure that each variable was measured in identical units for all companies. Some companies, for example, report accounting information referring to three-month or four-month periods for one or several years. In such cases, annual balance sheet figures were calculated as weighted averages of the multiple period values. In other cases firms reported numbers for a 12-month period, but the period did not coincide with the calendar year. The 1995 figures, for example, could refer to the period 1995-04-01 until 1996-03-31. In these cases, such 'deviations' were accounted for by adjusting the 'four quarter lag' (and thus the date at which the information is assumed to have been available) correspondingly.

From the set of balance sheet and income statement variables in Appendix B, a number of commonly used accounting ratios was constructed. We selected 17 ratios that were employed in frequently cited articles studying bankruptcy risk. See Altman [3], [4], [5], and [6], Frydman, Altman and Kao [17], Li [29], and Shumway [37]. Most of the accounting ratios are closely related liquidity measures, two are leverage ratios and the remainder are profitability ratios. In our empirical model, we employ three accounting ratios: earnings before interest, taxes, depreciation, and amortization (ebitda) over total assets (earnings ratio); total liabilities over total assets (debt ratio); and inventories over total sales (the inverse of inventory turnover). These three ratios were selected from the original list of 17 following a two-step procedure. First, the univariate relationship between the ratio and default risk was investigated. By visual inspection, ratios that displayed a clearly non-monotonic relation or lacked any correlation with default risk were deleted from the set of candidate explanatory variables. The left-hand side of Figure 3 illustrates these univariate relationships for the four variables we selected for the final model (shown in Table 2) - three accounting ratios and total sales (a proxy for firm size) - by comparing default rates (solid line) and the cumulative distributions of the variables (dotted

---

<sup>4</sup>Imputing the mean for missing values may lead to underestimation of standard errors. Little and Rubin [35] propose use of multiple imputations to overcome this problem in a situation where values are missing in a non-systematic manner. Since statistical significance is a minor concern with more than half a million observations, we have chosen not to apply their technique in the analysis.

Figure 3: Default rates and cumulative distribution functions for the accounting data.



line).<sup>5</sup> Default rates are calculated as averages over an interval of +/- 2,500 observations around the midpoint. There is a positive relationship between default risk and both the leverage ratio and the inverse of inventory turnover, while the figure suggests a negative relationship with both sales and the earnings ratio. We also checked if any significant differences in the average and median ratios existed between healthy and defaulting firms. Table 1 and Figure 3 contain some additional information on the distribution and the development over time for the financial ratios and non-payment and bank remarks. Table 1 shows that defaulting firms in general have lower earnings, lower sales, higher inventories and a higher level of indebtedness. The right-hand side panels of Figure 3 confirms this impression and suggests that these differences between (the median financial ratios of) healthy and defaulting firms are persistent, although possibly varying, over time. The median earnings of healthy enterprises, for example, are consistently more than twice as high as for defaulting ones. The difference in leverage ratio varies from approximately 15 percentage points in the mid-nineties to 25% in early 2000. On average, inventory turnover seems to be higher for defaulting firms, although there is some variation over time. Total sales differ in two respects between the two groups of businesses: they are strictly lower and vary more for defaulting firms than for healthy ones. This procedure led to the preliminary selection of seven candidate variables: the four that were selected for the final model of Table 2 and three measures of firm liquidity (cash over total assets, current assets over current liabilities, and accounts payable over sales).

In a second step, we studied their multivariate properties by estimating a number of permutations of the duration model. Neither of the three liquidity measures turned out to make a significant contribution in any of these permutations.

When selecting bank and legal remark variables we followed the same procedure. An intuitively reasonable starting point was to find remark events that (i) lead bank default as much as possible and (ii) are highly correlated with bank default. As it turned out, many remark variables are either contemporaneously correlated with default or lack any significant correlation with default behavior. An example of the first category is the start or completion of a company reconstruction. The most likely cause of this is the existence of a reporting lag. Tax related variables are typical examples of the second category. Of the remaining variables, many create a multicollinearity problem. For our final model, we selected two explanatory remark variables. One is a composite dummy of three events: a bankruptcy petition, the issuance of a court order - because of absence during the court hearing - to pay a debt, and the seizure of property. The

---

<sup>5</sup>We consider total sales (TS) to be a proxy for firm size. In principle, one can expect average TS to rise over time for two reasons. First, inflation will create a trend. Because inflation was low over the sample period, we chose not to deflate TS for the sake of simplicity. Second, one might expect average TS to grow over time if the average firm size increases over time. For our purposes, however, we need an absolute size measure, because larger firms generally have better access to credit and can be expected to be more robust to negative shocks. See for example Gertler and Gilchrist [18] [19].

other variable is "having a non-performing loan" with a bank, not necessarily with the current one. Table 1 shows that 10 respectively 20 % of the defaulting firms have a slow loan or a record of non-payment, in sharp contrast with the less than 1% among companies with performing loans.

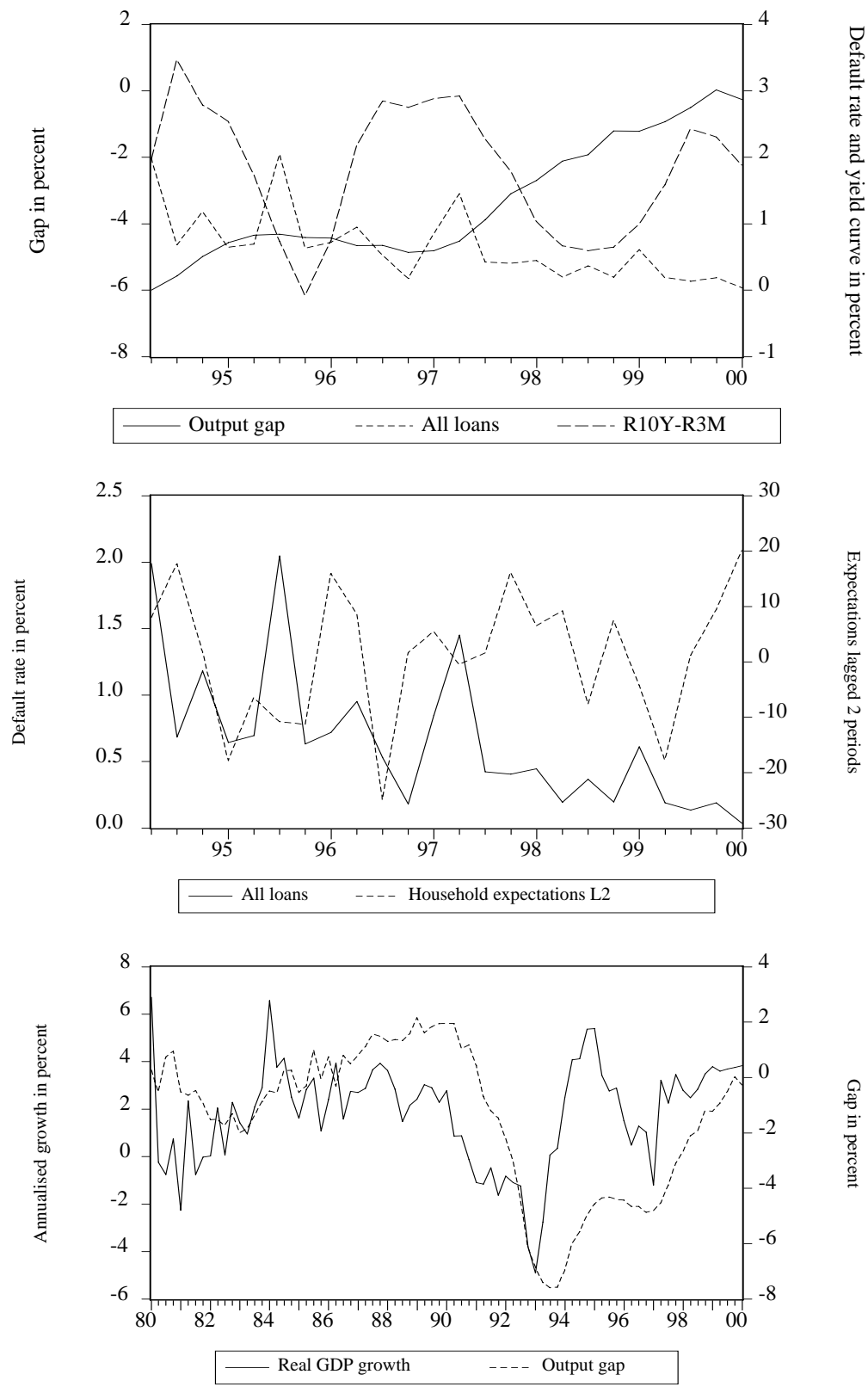
### 2.3 Macro data

The importance of macroeconomic effects for firm default risk and portfolio credit risk is so far a little explored topic in the empirical literature, at least in part due to a lack of suitable panel data on firms and their loan contracts. Intuitively, a stronger economy should drive down default risk for businesses. But what is the quantitative importance of the macroeconomy over and above idiosyncratic risk factors? And how should one capture these aggregate fluctuations? In this paper, our ambition is to contribute to these fields of research by combining the firm-specific bank data described above with a stylized description of the macroeconomy.

The last window in Figure 4 shows the developments of the growth rate in real GDP, in 1995 prices, and the output gap, given by the difference between actual and estimated potential GDP, for the period 1980 Q1 to 2000 Q2. The series for potential output is computed using an unobservable components method due to Apel and Jansson [7]. The deep recession in the beginning of the 1990's can be clearly seen in the figure, with negative growth figures (over 4 per cent per annum at most) and a negative output gap of more than 7 per cent. The general economic improvement of 1994-1996 is also evident. The first window shows how the yield curve and the output gap series are related to the average portfolio default rate. There is a strong downward trend in the default rate over the sample period, reflecting an improvement of the macroeconomic conditions. Finally, in the middle window we show the Swedish households' expectations of the future macroeconomic development, with a lag of 2 quarters, along with the average default rate of the portfolio. Worsening expectations are associated with increasing default rates.

A priori, we think that these three macroeconomic variables should have a measurable impact on the default risk of any given firm. The output gap functions as an indicator of demand conditions. Higher aggregate demand relative to production capacity can be expected to reduce default risk. Figure 4 seems, at large, consistent with this view, although there are some big spikes in the default rate that clearly have to be attributed to other variables. Research by Estrella and Hardouvelis [14] and Estrella and Mishkin [15] suggests that the yield curve can be an important indicator of future real activity; i.e., a positively sloping yield curve signals higher future economic activity and, reversely, a negatively sloped yield curve indicates an expectation of a fall in economic activity. Therefore, we expect that an increase in the spread between the short- and long-term interest rate is associated with decreasing default rates, since banks and firms will act upon this information: Banks will, for example, have stronger incentives

Figure 4: Macroeconomic variables and the average portfolio default rate.



to renegotiate loan terms or at least refrain from calling loans with firms that are at risk of bankruptcy given prospects of increased future demand. Likewise, firms may be more committed to avoid a default giving expectations of higher future economic activity. We use the difference between the nominal interest rates (annualized) on 10-year government bonds and 3-month treasury bills as a measure of the yield curve.

By similar reasoning, we expect that positive household expectations about increased future economic activity also reduce the default rate today. The index of household expectations about the future stance of the macroeconomy is taken from the survey data produced by the Swedish National Institute for Economic Research.<sup>6</sup> In the credit risk model, we will enter the series for the output gap and the household expectations with a lag of two quarters, since they are available for forecasting purposes with approximately that time delay. However, we will not lag the series for the yield curve spread, since this variable is accessible in real time.

### 3 Modeling default risk

In this section we outline the econometric model for the probability of default and estimate it using the bank's business loan data. So far we have treated a default as a discrete event, without being specific about the time-span within which this event might occur. In credit risk modeling the default event is usually defined in terms of a specific time-period, usually taken to be one year. Moreover, the application of using a time-specific definition is regularly complicated by the fact that one does not take account of the survival time of a loan, that is: one ignores any possible duration dependence. Duration dependence means that the time a borrower has managed to avoid default directly affects the risk of a default - over and above the effects from contemporaneous values of independent variables. In the absence of any duration dependence, there is no need for conditioning on the survival time of the borrower. As a consequence, the default variable can then simply be defined as a binary variable; the binary model approach will then also identify the parameters of a duration model for any arbitrary time-to-default horizon. However, if survival dependence does exist, then the model should account for it to avoid biased parameter estimates.<sup>7</sup> The model we outline in Section 3.1 below does consider survival dependence but allows for no-dependence as a special case.

---

<sup>6</sup>See the website of the National Institute for Economic Research for a description of this index: [www.konj.se/sidhuvud/inenglish/archives/archive/5.7d810b7d109c0650979800030671.html](http://www.konj.se/sidhuvud/inenglish/archives/archive/5.7d810b7d109c0650979800030671.html)

<sup>7</sup>Shumway [37] offers an investigation of the consequences of incorrectly assuming a constant default rate as implied by a binary model. He advocates the use of a duration model in order to avoid inconsistent and biased estimates.

### 3.1 The econometric model

Define the discrete random variable  $T$  to be the number of quarters it takes for a firm to default.  $\Pr[T = 1]$  is the probability that a firm defaults in the same quarter as credit was granted by the bank.  $\Pr[T = 2 | T > 1]$  is the conditional probability of default in the second quarter, given that the firm had not defaulted in the first quarter. In general, we are interested in the quantity  $\Pr[T = k | T > k - 1]$ ,  $k$  being a positive integer, which is usually referred to as the discrete time hazard rate. If  $\Pr[T = k | T > k - 1] = \Pr[T = k + l | T > k - 1 + l]$  for all positive, integer values of  $k$  and  $l$ , then default risk will be constant over time. For the data at hand we see no justification to stipulate that firm default risk does not change over time and hence we will estimate  $\Pr[T = k | T > k - 1]$  for  $k = 1, \dots, K$ , where  $K$  is the number of quarters in the sample period. For these data,  $K \leq 24$ .

Estimation of the hazard rate is in principle straightforward. The Maximum Likelihood estimator of  $\Pr[T = k | T > k - 1]$  is given by the number of firms that defaulted in the  $k$ -th quarter divided by the number of firms that had not defaulted prior to this quarter. There are however three technical problems involved.

The first problem concerns the question whether the definition of  $T$  is meaningful for this application. According to the definition, the outcome of  $T$  may be any integer greater than zero, and, as a consequence, the model will assign a probability of default greater than zero to any positive integer value of  $k$ . As a consequence, the model will even assign a positive probability to a default occurring after, say, 10,000 years, even though no loans are issued with such a duration.<sup>8</sup> One solution would be to redefine  $T$  such that its maximum value equals the intended duration (expressed in quarters) of the loan. However, we do not have any information on the intended length of the relationship between the bank and the borrower and the subject of our study is borrower risk - not loan risk. Therefore we find it reasonable to preserve the definition of  $T$  since the intention of (most) businesses is to avoid default eternally. In any case, the survival time of a business is generally not predetermined, making the choice of an alternative definition of  $T$  arbitrary.

The second problem concerns the issue of whether default is the only potential end state. In reality, most loans that leave the portfolio will not be observed to default. Rather, they exit because borrowers decide to pay off their debts. Thus, a loan can end up in one of two absorbing states; in other words, there are competing risks, c.f. Lagakos [27]. We will assume that the competing state is independent of the state of default. Hence, we can treat loans subject to such events as being censored in the last quarter that they were active.

The third problem concerns the relation between the sampling scheme and the population that we want to make inference about. Our wish is to be able to make inference about the

---

<sup>8</sup>It is, however, common for credit lines to be set up as open end contracts.

default risk of historical, current, or future borrowers. Therefore, we would ideally have data on historical, current, and future borrowers to estimate our model. In practice, we only have limited historical data, and hence assumptions are required.

The most important assumption is that of a constant inflow of borrowers into the portfolio (cf Lancaster [27]). Under this assumption one can use inflow data, i.e., entrants to the portfolio, to make inference about the population, or use a stock-sample, i.e., borrowers found to be in the portfolio at a specific point in time. Typically one combines both sampling schemes.

In our case, firms that entered the portfolio in 1994:2 and later are inflow data. Firms that were already in the portfolio before 1994:2 are stock data. The inflow data are directly relevant for the population and require no additional elaboration. The stock data, however, suffer from what is called length bias; borrowers are more likely to be in the stock if they are survivors, i.e., they have a small probability of default. Borrowers that default quickly are less likely to be in the stock. Consequently the average survival time of the stock tends to falsely give the impression that the time to default is long, simply because the quick defaulters are not in the data. To correct for this selection effect, three approaches exist, in principle, that all give consistent estimates under the assumption of a constant inflow (Lancaster [28]): (i) Assume that the pre-duration (i.e. the survival time of the borrower prior to the sampling date) and the post-duration (i.e. the survival time to default from the sampling date) are known. Let the full duration be the sum of the pre- and the post-duration. The first approach then uses the full-duration and weights each observation by its probability of being in the stock (which depends on the pre-duration). (ii) Solely use pre-duration data and estimate the model in the same fashion as in the case of inflow data only (when there is no length bias). (iii) This approach is similar to the preceding, but employs post-duration instead of pre-duration data. . For the borrowers that are in the portfolio in 1994:2 we do not know their pre-duration and therefore apply the third approach.

In Section 2 we discussed various factors that may affect default risk. These will be introduced in the model using the following notation:  $x$  refers to firm-specific factors, and  $z$  to factors specific to the general operating environment of all firms. We also make a distinction between variables that are constant - and those that vary over time. Hence, factors with a time index  $\tau$ , such as  $x(\tau)$  and  $z(\tau)$  are evaluated at calendar time  $\tau$ .<sup>9</sup> For example,  $x(\tau)$  will contain variables like loan size and various performance measures based on accounting data, as well as historical payment behavior of the firm, while  $x$  will consist of a number of loan type dummies. The purpose of having firm-specific variables in the model is to capture idiosyncratic risks. Business cycle effects are supposed to be captured by  $z(\tau)$ , which can include variables like the yield curve, the output gap, inflation, household expectations, and the unemployment rate. Environmental variables that do not vary over time ( $z$ ) are the duration dummies. The conditional probability

---

<sup>9</sup>The time-varying variables are assumed to be constant within each quarter.

of default, taking the heterogeneity of the borrowers into account, thus becomes:

$$\Pr [T = k \mid T > k - 1, x, z, x(\tau), z(\tau)], \quad (1)$$

The effects of  $x$  and  $z$  will be identified by cross-sectional variation in the probability of default while the effects of  $x(\tau)$  and  $z(\tau)$  are identified by cross-sectional variation in the default probability at *different calendar times*.

A common way to impose additional structure on the model is to assume a multiplicative relationship between the variables and the hazard rate suggested by Cox [12]. If we denote  $\Pr [T = k \mid T > k - 1, x, z, x(\tau), z(\tau)]$  by  $h(k)$  and  $\Pr [T = k \mid T > k - 1]$  by  $h_0(k)$  the duration model (1) takes the following intuitive form

$$h(k) = h_0(k) \exp(f[x, z, x(\tau), z(\tau); \beta, \gamma, \beta^\tau, \gamma^\tau]). \quad (2)$$

This model postulates a base-line hazard,  $h_0(k)$ , that is common to all loans and a multiplicative component that depends on both firm specific variables and common environmental variables. Here, the function  $f(\cdot)$  will be linear in  $\beta$  and  $\gamma$ , where the parameters  $\beta$  and  $\gamma$  pertain to the  $x$  and  $z$  variables. The superscript  $\tau$  indicates a parameter that is associated with a time-varying variable.

The data comprise a total of 54,603 firms and 69,249 lines of credit, which means that some firms are recorded with multiple credit lines.<sup>10</sup> Thus, there are 69,249 potential observations of  $T$ . We let  $k_i$  denote the  $i$ -th loan's duration, i.e., the number of quarters of survival prior to its default or, alternatively, its exit. Only 3,598 defaults were observed; the remaining durations of credit lines concerned either (non-defaulting) exits or survivals throughout the remainder of the 24 quarters. Let the censoring indicator,  $c_i$  be unity if the loan was observed to default and zero otherwise. The variables pertaining to the  $i$ -th loan will be indexed by  $i$ , and  $h_i(k)$  is the hazard rate evaluated for a loan with characteristics  $x_i$  and  $z_i$  and a time path of  $x_i(\tau - k : \tau)$  and  $z_i(\tau - k : \tau)$ . The maximum likelihood estimates of the parameters are obtained by maximizing

$$\ln L(h_0(k), \beta, \gamma, \beta^\tau, \gamma^\tau) = \sum_{i=1}^n \left( c_i \ln h_i(k) - \sum_{s=1}^{k_i-1} h_i(s) \right). \quad (3)$$

where  $n$  is the number of spells (lines of credit with multiple consecutive observations), i.e., 69,249.

---

<sup>10</sup>We stipulate that a bank relationship starts in the first quarter a firm is observed and continues until the first quarter for which the firm no longer is active. An interruption of one, or several quarters, is taken to indicate the end of a first relationship and the beginning of a second. A relationship ends only for one of three reasons; default (which is identified by a borrower-classification in rating class 15), repayment, presence of the firm in the last quarter of the sample period (2000:1). The latter two events imply that a relationship was incompletely observed. In those cases it is treated as a censored observation.

Our choice of explanatory variables is based on previously published work. However, the richness of the data has permitted us to explore an additional number of potentially important variables. Naturally, multicollinearity problems restrict one in choosing too opulent a set of variables. The guiding principles we applied when selecting variables are (in order of importance); (i) variables suggestions in earlier research, both theoretical and empirical, (ii) stability of the model - both in terms of predictions and parameter estimates, (iii) parsimony of the model, (iv) statistical significance of the estimated parameters, and (v) the fit of the model.

The way in which we selected variables from the set that has commonly been used in other work on firm default risk was described in Section 2.2. We verified that the stability criterion (ii) is satisfied by excluding, in turn, the following subsets; a 90 % fraction of the performing loans, loans after the second quarter of 1997, loans with missing values on at least one of the variables, and loans having values of the variables outside the 10 to 90 percentile range. As an additional check, we also included competing variables in the model. The third principle (iii) means that we avoid complicated transformations or interactions of variables, unless a substantial improvement in the fit could be achieved. The statistical significance of parameters (iv) is a standard requirement, although it is less useful as means to *select* variables when using very large data sets.<sup>11</sup> We use Goodman-Kruskal's  $\gamma$  as a *pseudo* -  $R^2$  measure of fit (v). This  $\gamma$ -measure can be interpreted as the degree to which the distribution of predicted probabilities of default for performing loans does not overlap the distribution of predicted probabilities of default for loans that actually defaulted.<sup>12</sup> The smaller the degree of overlap, the better the model discriminates between defaulted loans and non-defaulted ones, and hence, the better the predictive power of the model (See Fienberg [16] for details).

## 3.2 Estimation

Table 2 presents estimated coefficients and standard errors for three alternative model specifications.<sup>13</sup> Model 1 contains only balance sheet information from firms' financial statements as explanatory variables; Model 2 conditions on the variables in Model 1 plus two categories of payment remarks. Finally, the third model uses balance sheet information, payment remarks and macroeconomic variables. All models condition on a set of duration dummies.

---

<sup>11</sup>We neither include variables with t-ratios below two, unless the lack of significance of a variable is of interest *per se*.

<sup>12</sup>The  $\gamma$ -measure is based on the difference between concordant pairs (C) and discordant pairs (D) and calculated as:  $\Gamma = (C - D)/(C + D)$ .

<sup>13</sup>In general, there is no simple way of interpreting the coefficients in Table 2 in terms of expected duration - as one can in a linear regression model. However, if the hazard is independent of duration, then the expected duration increases by  $100 * (\exp(-\text{coefficient}) - 1)$  percent for a one unit increase in the variable. In the case of time-dependent hazards (here it increases over the life span of a loan), this effect needs to be multiplied by the increase in hazard parameter.

To compare the models in terms of explanatory power we have calculated two measures of model-fit, the above mentioned pseudo- $R^2$ , and an *aggregate* measure of fit.<sup>14</sup> The latter measure, labeled "aggregate"  $R^2$ , is given by the  $R^2$  in a regression of the quarterly average default rates on a constant and the predicted quarterly average default probabilities over the 24 quarters in our sample, where each predicted quarterly average default rate is calculated as the average over all estimated probabilities of default in a quarter, i.e.,  $\sum_{i=1}^{n_\tau} \hat{p}_{i,\tau}$ . The aggregate  $R^2$  measures to what extent the firm-level model of default captures economy-wide fluctuations in the rate of default - as opposed to firm-specific variation in the default rate that is captured by the model and picked up by the *Pseudo* -  $R^2$ .

Model specification 1 compares to the models used by Altman [3], [4] [5] [6], Frydman, Altman and Kao [17], Li [29], and Shumway [37], and can be treated as a benchmark, re-estimated with the data of this study. All financial ratios enter the model with the expected ratios, although the signs of the duration dummies indicates the possibly counterintuitive existence of a non-monotonic development of risk over the span of loans. A comparison of Models 1 and 2 in Table 2 shows that including legal and bank payment remarks yields a considerable increase in the model's fit: *Pseudo* -  $R^2$  rises from .294 to .400. At the aggregate level, the gain is smaller, and the fit improves merely from .258 to .334. Payment remarks could be considered to be information that is particular for Swedish banks - since many other countries, for instance the U.S., do not facilitate the institutionalized exchange of firms' track records. Table 2 makes it clear that expanding the benchmark model with payment remarks improves the fit while not drastically altering the relation between financial ratios and the likelihood of a default. The remarks thus capture a part of the variation in default probabilities that the financial ratios miss.

---

<sup>14</sup>We employ the square of Goodman-Kruskal's Gamma as pseudo  $R^2$  to make it directly comparable with the standard  $R^2$ .

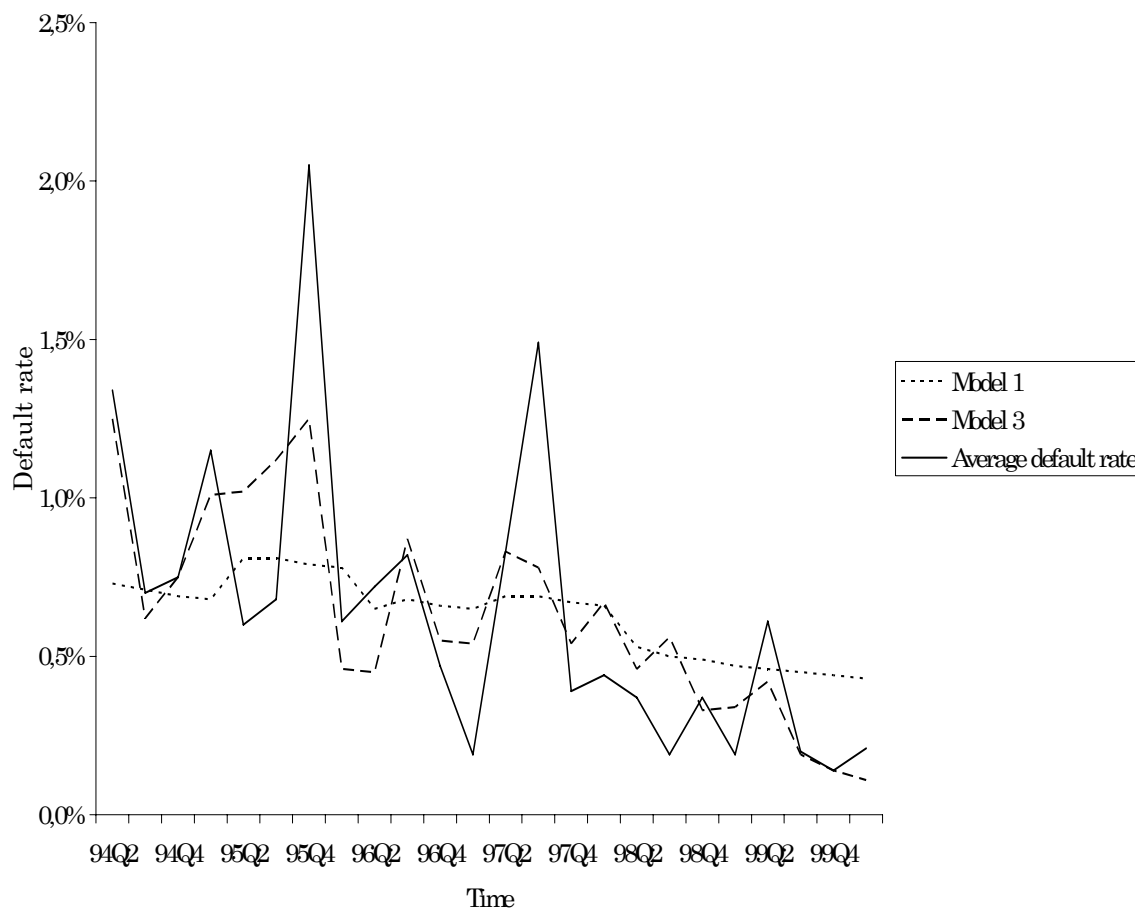
**Table 2:** The empirical credit risk models.

	Model 1		Model 2		Model 3	
	coeff.	s.e.	coeff.	s.e.	coeff.	s.e.
Duration						
1:st year	0	–	0	–	0	–
2:nd year	.256	.042	.317	.042	.396	.048
3:rd year	–.084	.051	.035	.051	.353	.056
4:th year	.141	.053	.254	.054	.688	.059
5:th year	–.435	.080	–.298	.080	.386	.096
6:th year	–.702	.115	–.568	.114	.849	.139
Credit type <sup>a</sup>						
Long-term	–1.189	.043	–.988	.044	–1.124	.044
Mix of short- and long-term	–.940	.070	–.736	.069	–.588	.069
Short-term	.270	.057	.223	.057	.073	.057
Accounting data <sup>c</sup>						
TS (mn SEK)	–.004	.001	–.003	.001	–.003	.001
EBITDA/TA	–1.053	.069	–.691	.068	–.768	.067
I/TS	.379	.070	.297	.069	.236	.069
TL/TA	3.597	.107	2.605	.103	2.367	.102
Remarks with credit bureau <sup>b</sup>						
Bank payment remark	–	–	.871	.063	.779	.063
Legal payment remark	–	–	2.637	.052	2.628	.052
Macroeconomic variables <sup>b</sup>						
Output-gap (lagged 2 quarters)	–	–	–	–	–.313	.018
Household expectations (first differences lagged 2 quarters)	–	–	–	–	–.024	.002
Yield curve, (10Y – 3M)	–	–	–	–	–.229	.022
<hr/>						
Pseudo- $R^2$	.294		.400		.437	
Aggregate- $R^2$	.258		.334		.593	

Notes: <sup>a</sup>variables taken to be constant over time, <sup>b</sup>variables taken to be time-varying, with quarterly variation, <sup>c</sup>variables taken to be time-varying, with yearly variation, although there is some variation between quarters due to the variation in reporting period that individual firms apply.

When introducing the output gap, a measure of the yield curve and households' expectations into the model (Table 2, third column), we find only a modest increase in *pseudo* –  $R^2$ , from about .40 to .44. This might lead one to incorrectly conclude that macro variables, although significant, have no important quantitative impact on default risk. The aggregate  $R^2$  measure, however, does display a large jump when going from Model 2 to Model 3. This can also be seen in Figure 5 that contains a plot of the average predicted default rates implied by Models 1 and 3 against the actual average default rate. Clearly, the model that includes macro explanatory variables does a much better job of capturing the variation in data than a model that is based on firm-specific information only. Apparently, while models based on idiosyncratic information

Figure 5: Actual and model predicted average default rates for the portfolio.



are able to rank loans or firms according to their riskiness, including macro-variables is essential to get the absolute level of default risk correctly estimated.

The accounting ratios and payment remarks are significant and enter the model with the expected signs in all three specifications. Both payment remarks and a rise in the leverage ratio and inventory ratio are associated with an increase in default risk while higher sales (a proxy for size) and earnings ratio reduce risk. Current real economic activity relative to trend (the output gap), (the slope of) the yield curve and households' expectations about the Swedish economy are significant indicators of the evolution of default risk over time. Among the macroeconomic variables, the output gap is most important for the default rate. The gap between actual and trend GDP varies from a low of  $-7$  per cent in the early part of the observation period to zero in the later part. This implies that the change in the output gap yields predicted default rates for the later part of the sample that are roughly 10 times ( $\exp(-.313 * -7)$ ) smaller than those for the early part of the observation period. The big spikes in the average default rate together

with the large fluctuations in macroeconomic aggregates, that occur during the years 1995 and 1997, are very helpful in identifying the effects of macroeconomic variables on default risk in the model.

The above considerations together make Model 3 our preferred model. The model has a number of interesting properties. First, default risk is almost monotonically increasing between years 2 and 6 - only year 5 displays a falling hazard. For example, the risk of default increases by roughly 50 percent in the second year of the loan, by 100 percent in the fourth year, and by 130 percent in the sixth year, all relative to the first year. All other equal, loans thus become riskier as they mature. We interpret this finding as reflecting the fact that firms that were recently granted credit are generally in good "shape", since they passed the screening process of the bank. As time passes, some of these firms become riskier and this average increase in riskiness is captured by the duration dummies. This is consistent with the results of Glennon and Nigro [20], who find that small-business defaults are time-dependent. Moreover, the results suggest that a logit or probit model, which implies a fixed default horizon, is inappropriate and would have resulted in inconsistent estimates.

Second, the risk of default is markedly higher for short-term loans than for long-term credit. Borrowers with a mix of short-term and long-term credit are significantly safer than short-term ones, but riskier than long-term ones. This finding is in line with the common idea that banks are more inclined to extend short-term loans to firms with a more uncertain future, while safer enterprises tend to have better access to loans with longer, fixed, maturities. Our findings support this view.

Having a legal payment remark registered in one or more of the preceding four quarters is a strong predictor for default. Of all defaulting firms, 30 percent received such a remark in the year before their default. Registration implies an increase in the risk of default by a factor  $\exp(2.628)$ , i.e., 12.85. If a firm receives a bank payment remark, default risk increases by approximately 20 percent. Here, one should keep in mind that getting a legal remark implies a firm has failed to satisfy at least some of its creditors during a longer time period, since it either submitted a bankruptcy petition or already went through a full court procedure. A bank remark is less common - only about 17 percent of all defaulted firms had one.

The accounting ratios also have considerable predictive power over and above payment remarks. Although the marginal effects of each variable will be different for each observation in a duration model, we can compare variables' marginal effects at the mean - both for performing and non-performing loans. When comparing the coefficients for sales, the earnings ratio, the inventory ratio and the leverage ratio across the three models one notices that they are somewhat reduced by the introduction of payment remarks in Model 2. But they remain almost unaffected by the inclusion of the macro variables in Model 3. While firm size, the earnings ratio, and the inventory turnover have a significant impact on default risk, the leverage ratio is

the financial ratio with by far the biggest impact on default risk - when evaluated at the mean. There's also substantially more variation in the leverage ratio than in the other financial ratios. If, for example, a firm's liability-ratio moves from 0.71, the mean for performing firms, to 0.84, the average for defaulting firms, then default risk increases by close to 200 percent.

## 4 Conclusions

Advances in computer technology, the aftermath of the Asian crisis in the 1990's and the subsequent work on the revision of the Basel Accord have spurred a renewed interest for research on default and credit risk. The last decade has brought about a range of techniques to model portfolio credit risk. In essence, four approaches have evolved: 'structural' models, that follow Merton's [36] approach, econometric factor risk models, that apply a 'bottom-up' method and compute default rates at either the individual firm level or at sub-portfolio level, 'top-down' actuarial models, that make no assumptions regarding causality, and non-parametric methods. Generally, the first three approaches build on similar mathematical structures and any discrepancies in their predictions generally result from differences in distributional assumptions or functional forms (Gordy [21]).

A number of empirical studies have considered the impact of macroeconomic conditions on default - or credit risk. Wilson [38] outlines some general principles behind McKinsey's proprietary portfolio credit risk model. Nickell, Perraudin and Varotto [32] use dummy indicators to integrate changes in the macroeconomic stance into a model of corporate bond rating transitions. Pesaran, Schuermann, Treutler and Weiner [33] employ a global macroeconometric model and link it to firm-specific returns in a Merton-type model. They make firm credit risk contingent on macroeconomic variables such as changes in equity indices, interest rates and output.

This paper contributes to the field by estimating a duration model for the survival time of firms' credit lines in a bank. Such a model can be used not only to analyze business defaults but also as an input in a portfolio credit risk model. The model includes macroeconomic explanatory variables and a set of more common firm-specific variables, such as accounting ratios and payment remarks. The width and length of the panel data allows us to separate firm-specific and macroeconomic effects by exploiting both the time-series and the cross-sectional dimension.

Our results provide a number of important insights. First, and most importantly, macroeconomic variables have significant explanatory power for firm default risk in addition to a number of common financial ratios. Both the output gap, the yield curve and households' expectations about the Swedish economy are quantitatively important indicators of the evolution of default risk over time. Their quantitative importance is also illustrated by the improved fit that we obtain. However, in the time-series dimension our sample is limited. Therefore our ability to characterize the impact of macroeconomic variables is likely to improve with an increased sample

period that also covers a full downturn of the economy. Second, a comparison of our model with a benchmark model of firm default risk that conditions only on firm-specific information shows that while the latter model can make a reasonably accurate ranking of firms according to their default risks, our model, by taking macro conditions into account, is also better able to capture the absolute level of default risk. Third, we find that firm default risk increases almost monotonically over the survival of their loans. This increasing hazard provides additional support for the supposed existence of duration dependence and is consistent with the results of Glennon and Nigro [20] This finding implies that idiosyncratic risk factors need to be complemented with information on survival time to obtain consistent default-risk estimates. It also means that a binary model specification is inappropriate. Finally, we find that the risk of default is markedly higher for short-term loans than for long-term credit. Borrowers with a mix of short-term and long-term credit are significantly safer than short-term ones, but riskier than long-term ones. This is in accordance with a common thought that banks are more inclined to extend short-term loans to firms with a more uncertain future, while safer enterprises tend to have better access to loans with longer, fixed, maturities.

## References

- [1] Acerbi, C., and D. Tasche, (2002), On the coherence of Expected Shortfall, *Journal of Banking and Finance*; 26 (7): 1487-1503
- [2] Allen, Linda and Anthony Saunders, (2004), Incorporating systemic influences into risk measurements: A survey of the literature, Forthcoming in: *Journal of Financial Services Research*.
- [3] Altman, Edward I., (1968), Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*; 23(4): 589-611.
- [4] Altman, Edward I., (1971), Railroad bankruptcy propensity, *Journal of Finance*; 26(2), pp. 333-345.
- [5] Altman, Edward I., (1973), Predicting Railroad bankruptcies in America, *Bell Journal of Economics*; 4(1): 184-211.
- [6] Altman, Edward I., (1984), The success of business failure prediction models, *Journal of Banking and Finance*; 8(2): 171-198.
- [7] Apel, Mikael and Per Jansson, (1999), "System estimates of potential output and the NAIRU", *Empirical Economics*; 24(3): 373-388.

- [8] Bangia, Anil, Francis X. Diebold, André Kronimus, Christian Schagen and Til Schuermann, (2002), Ratings migration and the business cycle, with application to credit portfolio stress testing, *Journal of Banking and Finance*; 26: 445–474.
- [9] Callem, Paul and Michael LaCour-Little, (2004), Risk-based capital requirements for mortgage loans, *Journal of Banking and Finance*; 28(3): 647-72
- [10] Carey, Mark, (1998), Credit risk in private debt portfolios, *Journal of Finance*; 53(4): 1363-1387.
- [11] Carey, Mark and Mark Hrycay, (2001), Parameterizing credit risk models with rating data, *Journal of Banking and Finance*; 25(1): 197-270.
- [12] Cox, David, (1972), Regression models and life tables, *Journal of the Royal Statistical Society*, Series B, 34: 187-220.
- [13] Estrella, Arturo, (2004), The cyclical behavior of optimal bank capital, *Journal of Banking and Finance*; 28(6): 1469-98.
- [14] Estrella, Arturo and Gikas A. Hardouvelis, (1991), The term structure as a predictor of real economic activity, *Journal of Finance*; 46(2): 555-576.
- [15] Estrella, Arturo and Frederic S. Mishkin, (1998), Predicting U.S. recessions: financial variables as leading indicators, *Review of Economics and Statistics*; 80(1): 45-61.
- [16] Fienberg, S.E., (1987), *The Analysis of Cross-Classified Categorical Data*, MIT Press, Cambridge, MA..
- [17] Frydman, Halina, Edward I. Altman, and Duen-Li Kao, (1985), Introducing recursive partitioning for financial classification: the case of financial distress, *Journal of Finance*; 40(1): 269-291.
- [18] Gertler, Mark, and Simon Gilchrist, (1993), The role of credit market imperfections in the transmission of monetary policy: arguments and evidence, *Scandinavian Journal of Economics* 95(1), pp. 43-64.
- [19] Gertler, Mark, and Simon Gilchrist, (1993), Monetary policy, business cycles and the behavior of small manufacturing firms, *Quarterly Journal of Economics* 109, pp. 309-340.
- [20] Glennon, Dennis, and Peter Nigro, (2003), *An Analysis of SBA Loan Defaults by Maturity Structure*, mimeo Office of the Comptroller of the Currency, Washington D.C..
- [21] Gordy Michael B., (2000a), A comparative anatomy of credit risk models, *Journal of Banking and Finance*; 24(1-2): 119-149.

- [22] Gordy, Michael B., (2003), A risk-factor model foundation for ratings-based bank capital rules, *Journal of Financial Intermediation*; 12(3): 199-232.
- [23] Hamerle, Alfred, Thilo Liebig and Daneil Rösch, (2002), *Credit risk factor modeling and the Basel II IRB approach*, mimeo, Deutsche Bundesbank.
- [24] Jacobson, Tor, and Kasper Roszbach, (2003), Bank lending policy, credit scoring and value-at-risk, *Journal of Banking and Finance*; 27(4): 615-633.
- [25] Jacobson, Tor, Jesper Lindé and Kasper Roszbach, (2006), *Internal ratings systems, implied credit risk and the consistency of banks' risk classification policies*, Forthcoming in : Journal of Banking and Finance.
- [26] Koyluoglu, H. Ugur, and Andrew Hickman, (1998), A generalized framework for credit risk portfolio models, forthcoming in *Risk*.
- [27] Lagakos, S.W., (1979), General right censoring and its impact on the analysis of survival data, *Biometrics*; 35: 135-156.
- [28] Lancaster, T., (1990), *The econometric analysis of transition data*, Cambridge University Press, Cambridge.
- [29] Li, Kai, (1999), Bayesian analysis of duration models: an application to Chapter 11 bankruptcy, *Economics Letters*, 63(3): 305-312.
- [30] Lucas, André, Pieter Klaassen, Peter Spreij and Stefan Straetmans, (2001), An analytic approach to credit risk of large corporate bond portfolios, *Journal of Banking and Finance* 25, (9) , pp. 1635-1664.
- [31] Lucas, André, Pieter Klaassen, Peter Spreij and Stefan Straetmans, (2002), Erratum to: "An analytic approach to credit risk of large corporate bond portfolios", *Journal of Banking and Finance* 26, (1) , pp. 201-202.
- [32] Nickell, Pamela, William Perraudin and Simone Varotto, Stability of rating transitions, *Journal of Banking and Finance*; 24(1-2): 203-227.
- [33] Pesaran, M. Hashem, Til Schuermann, Björn-Jakob Treutler, and Scott M. Weiner, (2005), *Macroeconomic dynamics and credit risk: a global perspective*, forthcoming in: Journal of Money, Credit and Banking.
- [34] Pesaran, M. Hashem, Til Schuermann, and Björn-Jakob Treutler, (2005), *Global business cycles and credit risk*, forthcoming in: Mark Carey and Rene M. Stultz (eds.), *The Risks of Financial Institutions*, NBER Volume, University of Chicago Press.

- [35] Little, R.J.A. and Rubin, D.B., (1987), *Statistical Analysis with Missing Data*, Wiley, New York.
- [36] Merton, Robert, (1974), On the pricing of corporate debt: the risk structure of interest rates, *Journal of Finance*; 29(2): 449-470.
- [37] Shumway, Tyler, (2001), Forecasting bankruptcy more accurately: a simple hazard model, *Journal of Business*; 74(1): 101-124.
- [38] Wilson, Thomas, (1997), Portfolio credit risk (I), *Risk*; 10(9): 111-117.
- [39] Wilson, Thomas, (1997), Portfolio credit risk (II), *Risk*; 10(10): 56-61.

## A Bank variables

EX\_DAT = Measurement date

BR\_K = 4 figure industry classification (by bank)

KO\_KU\_NR =

KU\_KAT =

BR\_GR\_K = 2 figure industry classification (by bank)

UT\_KRED = Amount of credit utilized

BE\_KRED = Granted credit

RI\_K = Rating class

KONK = Bankruptcy dummy

SAK\_BEL =

SEBRRKAP = Collateral 1

ENARRKAP = Collateral 2

KSKVPRGR = Dummy, 1 if short term credit is granted

KSKVBEKR = Amount of short term credit granted

LSLVPRGR = Dummy, 1 if long term credit is granted

LSLVBEKR = Amount of long term credit granted

S1S4PRGR = Dummy, 1 if mortgage is granted

S1S4BEKR = Amount of mortgage granted

GUARPRGR = Dummy, 1 if guarantee loan is granted

GUARBEKR = Amount of guarantee loan granted

MIXTPRGR = Dummy, 1 if other mixed credit is granted

MIXTBEKR = Amount of other mixed credit granted

## **B Credit bureau variables**

SAOMSTIL = Current Assets

SAKORTSK = Current Liabilities

SATILLG = Total Assets

LIKVID = Cash

VARULAG = Inventories

SALONGSK = Long Liabilities

LEVSKULD = Accounts Payable

SAANLTIL = Fixed Assets

SAEGETKA = Total Equity

OMSAETT = Total Sales

RESFOEAV = Earnings bef. Interest, Depreciation and Amortizations

ANTANS = No. employees

LOENER = Wages

AVSKRIVN = Depreciation

FININT = Financial income

FINKOST = Financial costs

EXTORDIN = Extraordinary costs

EXTORDKO = Extraordinary income

SKATT = Taxes

KUNDFORD = Accounts receivable

OVOMSTIL = Other liquid assets

SPAERRKO = Blocked accounts (e.g escrows)

GOODWILL = Goodwill

INVENT = Machinery etc

OBESRES = Untaxed reserves

AKTIEKAP = Nominal equity

OVREGBUN = Other Equity

SASKOEGE = Sum of taxes and equity (equals total assets)

SCBSNIKO = Statistics Sweden industry code

SCBSTKL = Statistics Sweden company size code

ORGNR = Company's 10 figure identification number

PANTER = Total of property pledges for non-mortgage loans)

ANSVAR = Total guaranties assumed for third party loans

SAFTGINT = Total of property pledges for mortgages in public register