WORKING PAPERS
RESEARCH DEPARTMENT

WORKING PAPER NO. 08-21
FIRM DEFAULT AND AGGREGATE FLUCTUATIONS

Tor Jacobson
Sveriges Riksbank

Rikard Kindell
Svenska Handelsbanken

Jesper Lindé
Sveriges Riksbank and CEPR

Kasper Roszbach
Sveriges Riksbank and
Visiting Scholar, Federal Reserve Bank of Philadelphia

September 19, 2008

RESEARCH DEPARTMENT, FEDERAL RESERVE BANK OF PHILADELPHIA
Ten Independence Mall, Philadelphia, PA 19106-1574 • www.philadelphiafed.org/research-and-data/

Firm Default and Aggregate Fluctuations*

Tor Jacobson Rikard Kindell Jesper Lindé Kasper Roszbach

19 September 2008

Abstract

This paper studies the relation between macroeconomic fluctuations and corporate defaults while conditioning on industry affiliation and an extensive set of firm-specific factors. Using a logit approach on a panel data set for all incorporated Swedish businesses over 1990-2002, we find strong evidence for a substantial and stable impact of aggregate fluctuations. Macroeffects differ across industries in an economically intuitive way. Out-of-sample evaluations show our approach is superior to both models that exclude macro information and best fitting naive forecasting models. While firm-specific factors are useful in ranking firms' relative riskiness, macroeconomic factors capture fluctuations in the absolute risk level.

Keywords: Default, default-risk model, business cycles, aggregate fluctuations, micro-data, logit, firm-specific variables, macroeconomic variables

JEL: C35, C41, C52, E44, G21, G33.

*Lindé: Sveriges Riksbank and CEPR. Jacobson and Roszbach: Sveriges Riksbank. Postal address: Research Division, Sveriges Riksbank, 103 37 Stockholm. Email: firstname.lastname@riksbank.se. Kindell: Svenska Handelsbanken, 106 70 Stockholm. This paper was partially written while Kasper Roszbach was a visiting scholar at the Federal Reserve Bank of Philadelphia. Discussions with and suggestions from Franklin Allen, Ed Altman, Mitch Berlin, Mark Carey, Bob Hunt, Wenli Li, Leonard Nakamura, Til Schuermann, Dragon Tang, and Kostas Tstsaronis have been very helpful in improving upon earlier drafts. We are indebted to Erik von Schedvin for providing outstanding research assistance, and grateful for comments from seminar participants at the EEA-ESEM meetings in Budapest, the National Bank of Hungary, the National Bank of Austria, EARIE, the 2008 BIS Rtf workshop, the Philadelphia Fed, Uppsala university and the 2008 ASSA meetings in New Orleans.

1 Introduction

The start-up and closing of businesses are two of the very fundamental events in economic life. Their rates of occurrence are associated with innovation and growth, and economic slowdowns, or even downturns, respectively. In spite of their fundamental importance, our understanding of these events is far from complete; little is known about the extent to which macroeconomic conditions influence the likelihood of start-ups and closings taking place. Recent economic events in the world economy, and the U.S. in particular, have once more confirmed that our understanding of the mechanisms by which and the extent to which the fortunes of businesses are influenced by broader economic conditions is still far from perfect. This is reflected in the difficulties economists experience in predicting the development of the economy in times of economic distress, in part due to a limited understanding of microeconomic responses to aggregate fluctuations.

The aim of this paper is to shed more light on the dynamics of business defaults and, in particular, the interaction between macroeconomic fluctuations on the one hand, and the firms' individual likelihood as well as the aggregate rate of default on the other hand. For this purpose we employ a new panel data set with detailed firm-level information on all incorporated Swedish businesses over the period 1990Q1-2002Q4. The panel contains more than 10 million data points and an average of over 200,000 firms per point in time. The length and width of this panel allow us to do several things that previous work has been unable to carry out. Among other things we can consider industry-specific effects of macroeconomic fluctuations and do extensive out-of sample testing of our models in both the cross-sectional and time-series dimensions.

Econometric studies of business defaults and bankruptcy risk started in the 1960s with work by Altman and coauthors (1968, 1971, 1973, 1984, 1985). These papers focused on explaining bankruptcies or defaults of publicly quoted businesses in a cross-sectional context with the help of a small set of firm-specific variables. Later work by Shumway (2001) explicitly models the time-dimension of defaults by means of a simple duration model. Jacobson, Lindé and Roszbach (2005) show how micro-econometric models of default can be incorporated into empirical macro models to capture the interaction between the real and financial part of the economy. Bharath and Shumway (2008) evaluate the out-of-sample accuracy of the Merton (1974) model and find that the distance-to-default measure is not a sufficient statistic for the probability of default.

Over time the average default frequency and individual default probabilities display substantial variation, in a way that suggests co-movement with macroeconomic and financial variables.

For aggregate default rates in the U.S. corporate bond market, the importance of macroeconomic conditions is well-documented by Blume, Keim and Patel (1991), Jonsson and Fridson (1996) and Helwege and Kleiman (1997). However, relatively little effort has been made to investigate the importance of macroeconomic fluctuations for business defaults, particularly for privately held companies. Recent work by Duffie, Saita and Wang (2007), Carling, Jacobson, Lindé and Roszbach (2007), Pesaran, Schuermann, Treutler and Weiner (2006) and Jacobson, Lindé and Roszbach (2005) provides some first empirical evidence that firm-specific factors alone cannot fully explain the variation in corporate default rates and the term structure of default probabilities. In these studies adding macroeconomic information contributes to explaining the likelihood of defaults. Using panel data on listed U.S. industrial firms, Duffie et al. (2007) find that macro variables, such as GDP growth and personal income growth, and firm size have no significant explanatory power for bankruptcy and default rates. Their model uses the distance-to-default as well as the trailing one-year stock return as firm-specific controls and the three-month T-bill rate and the one-year S&P 500 return as macro-financial covariates. Duffie et al. (2007) attain out-of-sample accuracy rates of over 80%.[1] Carling et al. (2007) document the significance of macroeconomic variables for Swedish loan defaults. Pesaran et al. (2006) focus on setting up a model that links credit losses to macroeconomic variables and use a large number of macro variables in a GVAR model to generate changes in Merton-model default probabilities for a hypothetical loan portfolio. Jacobson et al. (2005) follow a similar approach but link macro variables to a reduced-form model of loan defaults. Hackbarth, Miao and Morellec (2006) provide a first theoretical model of the mechanism through which macroeconomic conditions affect default risk. They argue that when cash flows depend on economic conditions, firms' optimal default thresholds will be affected by aggregate shocks. Hence aggregate shocks can trigger simultaneous defaults.

[1] Duffie et al. base their bankruptcy definition on the coding system in the Moody's database. This includes the following events: Bankruptcy, Bankruptcy Section 77, Chapter 10, Chapter 11, Chapter 7, and Prepackaged Chapter 11. A bankruptcy is also recorded when Compustat indicates Chapter 11 or Chapter 7. In some cases they record bankruptcy based on information from Bloomberg or other data sources. Defaults (see below) that eventually led to bankruptcies are sometimes not counted as bankruptcy exits if such a default was triggered earlier than the bankruptcy, for example by a missed debt payment. Defaults are defined as the occurrence of a bankruptcy as above or any of the following additional default types in the Moody's database: distressed exchange, dividend omission, grace-period default, indenture modified, missed interest payment, missed principal and interest payments, missed principal payment, payment moratorium, and suspension of payments. They also encompass any defaults recorded in Bloomberg and other data sources.

This paper is probably closest to Duffie et al. (2007), although we adopt a different methodology and use a very different data set. Our findings are consistent with those in Duffie et al. and most of the above-mentioned literature, but we are able to make several new contributions. First, we have access to an unusually large panel data set that includes all incorporated Swedish firms for a period covering several economic upturns and downturns. Hence our findings provide insights into the significance of aggregate fluctuations for both listed and privately held firms, the latter group typically being responsible for over half of GDP in developed economies. This feature is of importance because Merton-like models of default, which are based on stock price information, can only be applied to listed companies.[2] Second, we can make use of a nearly exhaustive set of firm-specific background variables. This allows us to look carefully at the importance of and interaction between firm-specific variables and macroeconomic information - an area that is more or less unexplored. Having access to a large set of firm specific controls makes it possible to eliminate any chance that the empirical significance of macro variables for default probabilities is (partially) an artifact of a shortage of firm-specific controls. Third, the length of our panel enables us to do extensive out-of-sample performance tests of our model.[3] Finally, the width of our panel permits us to investigate the relation between aggregate fluctuations and firm-defaults across industries. By isolating and comparing industry-specific effects of macro aggregates we get an additional measure of the robustness of the impact these macro variables have on business defaults.

We adopt a standard econometric specification and estimate logistic regressions on firm-level default risk.[4] In addition to an extensive set of financial statement variables and payment remarks, reflecting a firm's financial track record, we include four standard macroeconomic variables. The default risk models are estimated both at an economy-wide level and for industries at the one-digit level on the sub-sample covering $1990Q1 - 1999Q4$. For this period, we have 8,106,138 observations on roughly 250,000 firms. We assess the in-sample fit of the estimated models along with a thorough examination of the models' out-of-sample performance over the period $2000Q1 - 2002Q4$. Because the aggregate default frequency and macroeconomic aggre-

[2] In developed countries, privately held businesses' share of GDP typically exceeds 50 percent. Ayyagari, Beck and Demirguc-Kunt (2007) report that the share of SME's in GDP is between 50 percent and 60 percent in both Sweden and the United States. Hence, since a substantial share of large firms are privately-held, it is safe to infer that the share of all privately held firms is likely to be greater than the share of SME's alone. See also Kobe (2007).

[3] Duffie et al. (2007) do out-of sample tests for the subset of listed industrial firms.

[4] See Altman and Saunders (1997) for references.

gates displayed a substantial amount of volatility during the 1990s, we consider the out-of-sample tests as an important step in our assessment of the extent to which our model can be viewed as causal. Out-of-sample accuracy for the economy-wide and the industry models lends support for our hypothesis that aggregate fluctuations are important for understanding default behavior at the firm level, over and above the effect of an extensive set of firm-specific factors.

Our main findings are as follows. First, we find that macroeconomic variables are important for explaining the time-varying likelihood of default. Firm-specific variables are very useful for ranking firms according to their relative riskiness, but macroeconomic variables are of crucial importance for explaining variation in the level of default risk over time. Second, our analysis also suggests that considering only macro variables while ignoring relevant firm-specific information leads to a substantial loss of out-of-sample accuracy. Third, the quantitative effects of aggregate fluctuations in the industry-specific models are such that they support the notion that the macro factors we consider have truly causal effects. For example, demand and interest rate conditions have a particularly strong impact on the construction and real-estate sectors while the dependence of the agricultural sector on the macroeconomic stance is relatively weak. Fourth, we show that models estimated on cross-sectional data suffer from a substantial parameter instability. Fifth, we document that the estimated default risk models perform very well out-of-sample, along the cross-sectional and the time-series dimensions as well as at the aggregate and the industry levels. As may be expected, industry-specific models typically have a clear edge over a single economy-wide model when evaluated at the industry level. However, when evaluated at an economy-wide level, this advantage more or less vanishes. By and large, we think these findings are of great interest, since they imply that even economic outcomes that are generated in a period with extreme aggregate fluctuations, such as the Swedish banking crisis in the early 1990s, can be captured by a default risk model with constant parameters over time.

The remainder of this paper is structured as follows. In the next section, we present our micro and macro data sets. The logistic regression results are presented in Section 3 for two versions of the model, one where only firm-specific variables are included and another where the model is extended with macroeconomic variables. We also compare the industry-specific models with the estimation results of an economy-wide model, and make an assessment of the in-sample fit of the estimated models. In Section 4, we undertake a thorough out-of-sample investigation of the estimated models along three dimensions: *i*) the fit of the models in terms of adjusted R^2, *ii*) the root mean squared prediction errors and *iii*) the accuracy of the default risk ranking.

The former two measures are studied at the industry and the economy-wide level, while the latter criterion is an assessment of the microeconomic relevance. Finally, Section 5 concludes.

2 Data

In this section we will, at some length, discuss the very large micro data set at hand. We also briefly cover the macro data.

2.1 Micro data

The firm data set is a panel consisting of $10,720,386$ quarterly observations on the stock of Swedish *aktiebolag*, or firms, between January 1, 1990, and December 31, 2002, hence covering a period of 13 years. *Aktiebolag* are by approximation the Swedish equivalent of US corporations and UK limited liability businesses. Swedish law requires every *aktiebolag* to have at least SEK $100,000$ (approximately US$ $15,000$) of equity to be eligible for registration at *Bolagsverket*, the Swedish Companies Registration Office (SCRO). Swedish corporations are also required to submit an annual report to the SCRO. Although the corporate sector in Sweden also includes small firms such as general partnerships, limited partnerships, and sole proprietors, we will not include them in the analyses for a number of reasons. First, these firms do not submit yearly financial statements and hence require model specifications that do not involve financial ratios. Second, the incorporated firms that we model account for an overwhelmingly large fraction of Swedish bank loans to firms. Third, corporations display a more pronounced cyclical variation in default risk; see Jacobson and Lindé (2000).

The firm data have been obtained from Upplysningscentralen AB (UC), the leading credit bureau in Sweden, independently operated but jointly owned by the Swedish banks. The UC data come from two general sources. The first concerns balance-sheet and income-statement data from the firms' compulsory annual reports submitted to the SCRO. These annual report data cover the period January 1, 1989, to December 31, 2002, and the format is in accordance with European Union standards.

The second information source is atypical in the default literature and somewhat unique for Sweden. The credit bureau systematically collects information about events related to firms' payment behavior from all relevant sources, e.g., the Swedish retail banks, the Swedish tax authorities, and the institutions that deal with the legal formalities in firms´ bankruptcy processes.[5] The credit bureau thus has a register of more than 60 different payment remarks concerning foremost credit and tax-related events but also records of various steps in the legal procedures leading up to formal bankruptcy. The information in the register involves a flag for the occurrence of an event in the form of a date and the amount of due payment (if applicable). Some examples of registered events are delays in tax payments, the repossession of delivered goods, the seizure of property, the restructuring of loans, and actual bankruptcy. The storage and usage of payment remarks are regulated by the Credit Information Act and the Personal Data Act are overseen by the Swedish Data Inspection Board. Payment remarks turn out to be powerful predictors of default in practice. With a record of remarks individuals will usually not be granted any new bank loans, and businesses can find it very difficult to open new lines of credit.

For this study, we define the population of existing firms in quarter t as the firms that have issued a financial statement covering that quarter and are classified as "active," i.e., the firm has reported total sales and total assets in excess of $1,000$ SEK (roughly US$ 150). However, since there are firms that neglect to fulfil their reporting obligation, a behavior typically associated with distress, we would miss an important segment of firms by only considering those that submit annual reports regularly. Hence we will add the firms that, according to the data set with payment remarks, are classified as defaulted firms in quarter t. Many firms that default choose not to submit their compulsory annual reports in that year or even for a number of years prior to default. Hence, the only records of their existence that we have come from the payment remark registers. We adopt the following definition of a default: a firm has default status if any of the following events has occurred: the firm is declared legally bankrupt, has suspended payments, has negotiated a debt composition settlement, is undergoing a re-construction, or is distraint without assets. More details on the construction of the default variable are provided in the Appendix.

In Table 1, we report the means and standard deviations for a set of accounting ratios, payment remarks, and a variable that measures the average elapsed time since the last issued financial report. The table distinguishes between defaulted and non-defaulted firms, at the aggregate as well as the industry level, for the in-sample period $1990Q1 - 1999Q4$, that is, the sub-sample period for which we will specify and estimate all subsequent models. The out-of-

[5] District courts, the Swedish Enforcement Authority, the Swedish Companies Registration Office, and debt collection firms, among others.

sample period, $2000Q1 - 2002Q4$, is saved to allow for extensive model-evaluation exercises. Analyses of industry effects will be conducted at the one-digit level to ensure sufficiently many default observations in each industry in both the cross-sectional and the time series dimensions. The ten industries are; agriculture, manufacturing, construction, retail, hotel and restaurant, transportation, banking, finance and insurance, real-estate, consulting and rental, and finally a residual industry labelled "not classified".

Because of the varying availability of data, the statistics in Table 1 are calculated based on slightly different numbers of observations for the variables in a given industry. Dealing with micro data sets of this size invariably involves dealing with outliers. As indicated by the large standard errors in Panel A of Table 1, showing non-winsorized data, there are some accounting data observations that clearly are severe outliers. These observations would distort the estimation results if they were to be included in the logit model and therefore, we have winsorized the top and bottom 1 percent observations for the accounting variables in each industry.[6] Given the large number of observations in our data set, this approach is practically more or less equivalent to simply deleting 1 percent of the observations that have accounting data that fall outside a certain region. Note that we choose to winsorize the observations in each industry separately, rather than at the aggregate level, thereby implicitly allowing for dispersion and different means in different industries. Panel B of Table 1 shows the descriptive statistics for the truncated micro data set.[7]

In deciding on the set of financial ratios in Table 1, we evaluated a larger number of frequently used ratios in often-cited articles on bankruptcy risk and the balance-sheet channel.[8] The six financial ratios reported show the strongest correlations with our definition and measure of default: earnings before interest, depreciation, taxes and amortization (EBITDA) over total assets (TA) (*earnings ratio*); interest payments (IP) over the sum of interest payments and earnings before interest, depreciation, taxes and amortization (*interest coverage ratio*); total liabilities (TL) over total assets (*leverage ratio*); total liabilities over total sales (TS) (*debt*

[6] Winsorization is quite common in the literature using financial ratios to avoid outliers that are created by near-zero denominators. Shumway (2001) winsorizes the top and bottom 1 percent of all observations. It should be emphasized that the results are robust to varying the winsorization rate between 0.5 and 2 percent.

[7] From Table 1, comparison of the descriptive statistics for the unwinsorized data makes it clear that defaulted firms are disproportionally more affected when winsorizing all observations jointly. Since the PAYREMARK, TAXARREARS, PAYDIV and TTLFS are dummy variables that are unaffected by choice of winsorization procedure, a joint one could lead to underestimation of the importance of the accounting data variables in the default risk model relative to these dummy variables. To check the robustness of our chosen approach, we used an alternative approach where we truncated the healthy and defaulted firms separately. As expected, the estimation results of the default-risk model with this alternative winsorization suggest a somewhat larger role for the accounting ratios, but the overall picture remains the same.

[8] See Altman (1969, 1971, 1973, 1984), Carling et al. (2007), Frydman, Altman and Kao (1985), and Shumway (2001).

ratio); liquid assets (LA) in relation to total liabilities (*quick ratio*); and inventories (I) over total sales (*inventory turnover ratio*).[9] These six ratios were selected following a two-step procedure. First, the univariate relationship between the ratio and default risk was investigated. By visual inspection, ratios that lacked any correlation with default risk were eliminated from the set of candidate explanatory variables. Figure 1 illustrates this for the six selected ratios by comparing default rates (solid line) and the cumulative distributions of the variables (dotted line) for all observations in the panel $1990Q1 - 1999Q4$. The default rate for a given observation of a ratio is calculated as an average over the interval of +/- 5000 adjacent observations in the empirical distribution of the ratio at hand. Given the density of the observations, there is a positive relationship between default and the leverage, interest coverage and turnover ratios, while the figure suggests a negative relationship for both the debt and the liquidity ratios. The diagrams in Figure 1 suggest that the relationship between default and the earnings ratio, total liability over total sales ratio and interest costs over the sum of interest costs and earnings are non-linear. For instance, for the interest coverage variable, this relationship is perhaps what one would have expected. The ratio can turn highly negative if earnings are negative and slightly larger than interest payments in absolute value, which is intuitively associated with an increased risk of default. On the other tack, large interest payments and low earnings will also make this ratio large, which is likewise associated with an increased default risk. Similar reasoning can be be applied to the other ratios. What is important to note is that this non-linear feature of some financial ratios does not imply that these variables are uninformative for default risk in the empirical models, even when entered linearly in the logit model. The reason for this is that the co-variation between these financial ratios in the cross section is substantial, which makes each of these variables contribute to predicting default risk in the joint empirical model.[10] Taking these insights into account, Figure 1 confirms the picture emerging from Table 1: there is a clear difference between healthy and defaulted firms for these variables. In the accounting data, we also have information on whether a firm has paid out dividends to shareholders or not. We therefore include this information as a dummy variable (PAYDIV) in the model, taking a value of one if the firm paid out dividends and zero otherwise.

As mentioned previously, some firms classified as active or defaulted did not submit a financial

[9] It should be noted that the level of debt, in addition to the leverage ratio ($\mathrm{TL}_{i,t}/\mathrm{TA}_{i,t}$) for firm i in period t, contains predictive power for default. We therefore decided to include $\mathrm{TL}_{i,t}$ as a separate variable, but scaled it with average total sales in period t to obtain a stationary ratio. Thus, the debt-to-sales ratio is defined as $\mathrm{TL}_{i,t}/\mathrm{TS}_t$,where TS_t denotes average total sales in period t.

[10] For instance, taking the square of the interest coverage ratio, which, judging by Figure 1, would seem appropriate in a single-variable analysis, reduces the explanatory power of this variable in the multivariate model.

report in every period, leading to a missing observation problem. Rather than excluding such firms from the sample, we replace missing values by imputing the panel mean for the joint set of defaulted/non-defaulted firms. In order to capture the relationship between the failure to a financial statement and subsequent default, we also include a dummy variable, denoted TTLFS, which equals unity if a firm has not issued a financial statement one and a half years prior to the current quarter, and zero otherwise.[11] By comparing defaulting and healthy firms in Table 1 we see that this mechanism is at work in the panel.

For the remark variables, we employ the same approach as in Carling et al. (2007) and use simple dummy variables by setting them to unity if certain remarks existed for the firm during the year prior to quarter t, and 0 otherwise. An intuitively reasonable starting point was to find remark events that (i) lead default in time as much as possible and (ii) are highly correlated with default. As it turns out, many payment remark variables are either contemporaneously correlated with default or lack a significant correlation with default behavior. For our final model, we constructed the PAYREMARK variable as a composite dummy of four events: "a bankruptcy petition," "the issuance of a court order - because of absence during the court hearing - to pay a debt," "the seizure of property," and "having a non-performing loan." The TAXARREARS variable reflects whether the firm is in various tax arrears. It should be emphasized, although it is evident from Panel B in Table 1, that the constructed payment remark variables that we consider do not automatically imply a subsequent default incident, so there are no tautological issues involved in using these variables to predict default events.

There is some, but not very much, variation in the average financial ratios and payment remark variables across industries, and in general the differences between defaulted and non-defaulted firms display similar patterns in all industries. So, for example, in Table 1, panel B, we see that the shares of defaulted firms that have received payment remarks are around 0.15 and 0.45, respectively, whereas corresponding shares for non-defaulted firms are 0.00 and 0.03. The hotel and restaurant industry is the outlier. Hence, these firms have the lowest earnings ratios, largest debt ratios, greatest occurrences of payment remarks and least of dividend payments,

[11] Three things worth noting in connection with the definition of TTLFS. First, this information is assumed to be available with a 6-quarter time lag, since financial statements for year τ are typically available in the third quarter of year $\tau + 1$. By letting this dummy variable equal unity with a 6-quarter time lag in relevant cases, we account for the real-time delay. Second, given the way we define the population of existing firms, firms that recently registered and entered into the panel would automatically be assigned TTLFS = 1 in the third quarter of their existence, since they have not, of course, issued any financial statement prior to entering. For these new firms, TTLFS has been set to 0 and the accounting data variables have been taken from their first yearly balance sheet and income statement. Third, for defaulting firms that are in the panel but on no occassion submitted an annual report, we also set TTLFS equal to 0. This is the case for $49,202$ out of $123,023$ defaulting firms in the panel. So, although TTLFS turns out to be very important in the default-risk model, by construction the importance of this variable is down-played rather than exaggerated.

and as a consequence, the largest default rate over all.

2.2 Macro data

In this paper, we will make use of the same macrodata set as in earlier work, (see Jacobson et al. (2005)) and consider the output gap (i.e., the deviation of GDP from its trend value), the yearly inflation rate (measured as the fourth difference of the GDP deflator), the REPO nominal interest rate (a short-term interest rate, set by the Riksbank), and the real exchange rate.[12] The output-gap series is by construction a detrended variable. Since the real exchange rate is also characterized by a strong trend during the sample period, this variable is detrended as well in order to allow for a closer relationship with firm default. Since Sweden is a small and open economy with a large foreign trade sector, one should consider the importance of conditioning on foreign variables in the empirical analyses as well. Our results suggest that while foreign variables are an important source of fluctuations in Swedish macro variables (see Lindé, 2002), it is not necessary to condition directly on foreign variables in the default risk models if the above-mentioned domestic variables are included. The aggregate time series are depicted in Figure 2.[13]

In the output-gap series in Figure 2 there is clear evidence of the deep recession in the beginning of the 1990s with a negative output gap of more than 4 percent in 1993. The general economic improvement of 1994-1996 is also evident. This can also be seen in the inflation and interest rate series that both peak in the early 90s and then come down in the recovery phase.

3 The default-risk models: Estimation and in-sample fit

In this section, we examine if default risk at the firm level is affected by aggregate fluctuations over and above the set of firm-specific information that we have at our disposal for all incorporated firms.

We study the in-sample gains of estimating separate models for each industry and assess the role of aggregate fluctuations for improving the models' fit. The in-sample period is chosen to be $1990Q1 - 1999Q4$.

[12] The real exchange rate is measured as the nominal TCW-weighted (TCW= trade competitive weights) exchange rate times the TCW-weighted foreign price level (CPI deflators) divided by the domestic CPI deflator.

[13] The macro data set has been adopted from Lindé (2002) and is based on an estimated vector autoregressive model (VAR) for Sweden and the period $1986Q3 - 2002Q4$. The trends for the variables are estimated by means of a dynamic simulation of the estimated VAR under the assumption of no shocks hitting the equations. The detrended variables are then computed as actual values minus the trend values. It should be noted, however, that using HP-filtered data for output and the real exchange rate produces very similar results to those reported.

3.1 The default-risk models

The reduced-form statistical model that we employ for estimating probabilities of default for all Swedish incorporated firms is similar to the logit approach used in Shumway (2001), and in Jacobson, Lindé and Roszbach (2005). The specification includes both firm-specific ($x_{i,t}$) and macroeconomic explanatory variables (z_t). Using a reduced-form model both avoids the problem that the Merton (1974) model cannot be implemented for privately held companies without very strong assumptions and enables us to use a unified approach for all businesses, both privately and publicly held. Our approach is consistent with the theoretical ideas in Hackbarth, Miao and Morellec (2007), who argue that aggregate shocks can trigger simultaneous defaults. Thus we propose to estimate the following model:

$$\begin{aligned} y_{i,t} &= x_{i,t}\beta + z_t\gamma + \varepsilon_{i,t}, \\ & \text{where} \\ y_{i,t} &= \begin{cases} 1 \text{ if } x_{i,t}\beta + z_t\gamma + \varepsilon_{i,t} \geq \ 0 \text{ (firm defaults)} \\ 0 \text{ if } x_{i,t}\beta + z_t\gamma + \varepsilon_{i,t} < \ 0 \text{ (firm stays in business)} \end{cases}, \end{aligned}$$

under the assumption that the vector of firm-specific regressors (i.e. $x_{i,t}$) and the macroeconomic variables we consider (collected in the vector z_t) are stochastically independent w.r.t. the error term $\varepsilon_{i,t}$. We also make the assumption that the errors are independent between both firms and time points, i.e., $f(\varepsilon_{i,t}, \varepsilon_{j,t}) = f(\varepsilon_{i,t}) f(\varepsilon_{j,t})$ for $i \neq j$ and $f(\varepsilon_{i,t}, \varepsilon_{i,t+p}) = f(\varepsilon_{i,t}) f(\varepsilon_{i,t+p})$ for $p \neq 0$.

We use standard macro variables in the model: a measure of the output gap, the domestic yearly inflation rate, the REPO rate (a short-term nominal interest rate controlled by the Swedish central bank), and the real exchange rate. These variables are depicted in Figure 2. Although the literature does not offer a strong theoretical basis for selecting macro variables, we think a priori that these variables could credibly have measurable impact on the default risk of any given firm. The output gap is intended as an indicator of demand conditions, i.e., increased demand in the economy is expected to reduce default risk. We also include the nominal interest rate in z_t because credit conditions facing firms, in particular firms in distress, are likely to be tightly linked to levels of the interest rate. In addition, the nominal interest rate displayed considerable variation during the recession in the early 1990s but has since come down substantially, after the adoption of an inflation target in Sweden.[14] Given the fact that the exports-to-GDP-ratio in Sweden is around 0.40, the real exchange rate is also a potentially important variable, a

[14] The REPO rate was extremely high in the third quarter of 1992 due to the Riksbank having raised the so-called marginal interest rate to 500 percent, unexpectedly and temporarily, in an attempt to defend the fixed

deprecation rendering improved competitiveness to Swedish firms. The inflation rate may also be important for firms' pricing decisions; higher inflation rates are associated with less certainty about correct relative prices and may thus lead to potentially higher default risks.

3.2 Estimation results

To document how aggregate variables contribute to the default risk models, we present estimation results for two specifications: one with and one without macroeconomic variables. Moreover, results are presented for ten industry-specific models, as well as an economy-wide model (all firms in all industries jointly modelled), and also results achieved by aggregating across the industry models using industry size as weights.

Table 2 contains estimation results for a model with firm-specific determinants of default risk only (i.e., the six financial ratios augmented with the dummy variables PAYDIV, TTLFS, PAYREMARK, and TAXARREARS), while Table 3 shows results with the macroeconomic variables added.

Since the firms' annual financial reports are typically submitted with a significant time lag, it cannot in general be assumed that accounting data for year τ are available during, or even at the end of, year τ and enable forecasted default risks for the year $\tau+1$. To account for this, we have lagged all accounting data by 4 quarters in the estimations. For most firms, which report balance-sheet and income-statement data over calendar years, this means that data for year τ are assumed to have been available in the first quarter of year $\tau+1$. It should be emphasized that our decision to lag the accounting data 4 quarters in the estimation in order to make the model "operational" in real time has minor implications for the estimated coefficients. When re-estimating the model using contemporaneous data instead, the estimation results were found to be very similar to the ones reported in Tables 2 and 3.[15]

exchange rate. If the REPO rate is not adjusted for this exceptional event, the estimation procedure would lead to underestimation of the importance of financial costs for default behavior. We therefore decided to adjust the REPO rate series in the third quarter of 1992. The estimated dummy coefficient in the VAR that we used to compute the output gap and the real exchange rate gap equals 28.2 in the REPO-rate equation. On the basis of this, we have adjusted the REPO rate for this quarter to equal 9.8 percent instead of 38 percent.

[15] In addition to the coefficients reported in Tables 2 and 3, three more variables were included (but not reported). First, an industry-specific intercept. Second, since the bankcruptcy rate is systematically lower in the third quarter (most likely due to Swedish courts' summer holiday period in July-August), a seasonal dummy is included to capture this phenonemon. Third, because no data on the payment records of firms (i.e., the dummy variables PAYREMARK and TAXARREARS) exist prior to 1992$Q3$ for legal storage reasons, the models also include one additional variable common to all i firms that is constructed to be an estimate of the average value of the sum of the payment record variables PAYREMARK and TAXARREARS for the quarters 1990$Q1$-1992$Q2$. This variable was constructed by estimating a logit model for the event of either of the dummy variables PAYREMARK and TAXARREARS taking on the value 0 or 1 for the period 1992$Q3$-1999$Q2$, using all the variables in the model in Table 3 as regressors (except PAYREMARK and TAXARREARS, of course). The imputed average value for this variable for the period 1990$Q1$-1992$Q2$ (after 1992$Q2$, it is set to nil) was then

The results in Table 2 show that the firm-specific information we consider is indeed important for explaining default behavior in both the industry-specific models and in the economy-wide model. In particular, the indicator variable TTLFS (which takes a value of 1 if a firm has not filed an annual report on time, and 0 otherwise) and the variables for remarks on firms' payment records are very powerful predictors of default. Among the financial ratios we find the leverage ratio and the debt ratios TL/TA and TL/TS as well as the earnings ratio to be quite useful.[16] However, the turnover ratio, the quick ratio, and the interest coverage ratio appear to be less important. Moreover, the roles played by financial ratios in the various industry models differ substantially; while accounting data are less important in the financial services (bank, finance and insurance) sector, it is more important in the manufacturing industry. In the hotel and restaurant sector, we find that the I/TS coefficient is large, whereas it is zero, or even negative, in the agriculture and construction industries, respectively. The coefficients for the payment remarks and the indicator variable TTLFS are quite similar across industries. So to the extent that these variables are the more important ones for explaining firm default behavior, there is no clear gain at the firm-specific level from conditioning on industry. Finally, a reassuring feature of the results in Tables 2 and 3 is that the coefficients for the firm-specific variables do not change substantially when the model is augmented with the macroeconomic variables. In particular, coefficients for the financial ratios in Table 2 are in general very similar to the ones in Table 3.

Turning to the estimation results presented in Table 3 for models with the macroeconomic variables included, we find that all coefficients are significant in the economy-wide model, with the exception of inflation, and have the expected signs.[17] The notion of conditioning on macroeconomic variables in default risk modeling is given further support by the industry-specific

constructed as the average estimated probability for each firm and period, i.e., $\mathrm{RD}_t = \frac{1}{N_t}\sum_i \hat{p}_{i,t}$ where $\hat{p}_{i,t}$ denotes the estimated probability for firm i in period t to have either PAYREMARK or TAXARREARS greater than zero, and N_t denotes the number of firms in period t. The largest gain in including this variable is that presumably the effects of macroeconomic variables in Table 3 are somewhat more accurately captured. For the coefficients of the firm-specific variables this imputation is of little consequence.

[16] Regarding the importance of the accounting data in the model, we would like to emphasize the following. Firms issue annual financial statements, which we transform into quarterly observations by assuming that the variables for a firm remain constant over the quarters in a given reporting period. By defining a default event at quarterly frequency, our transformation procedure could potentially lead to underestimation of the importance of the financial statement variables in the default-risk model. As a robustness check we estimated the default-risk models on an annual frequency instead and found that the coefficients for the accounting variables are quite similar for either frequency specification. In the economy-wide model, only the coefficients for the earnings ratio, EBITDA/TA, and the leverage ratio, TL/TA, were found to be slightly lower/higher (−1.13/0.59 instead of −0.95/0.49, respectively). The coefficients for the other accounting variables were found to be very similar. As for the indicator variables, TAXARREARS and TTLFS were found to be somewhat smaller in the annual model (2.30/3.07 instead of 2.57/3.67, respectively), but the coefficients for the other dummy variables PAYREMARK and PAYDIV were basically unaffected.

[17] Note that a larger value for the real exchange rate implies a depreciation and therefore a negative estimated coefficient for this variable implies that a depreciation on average reduces the risk of default at a given point in time.

model results. Table 3 shows that the impact of the macroeconomic factors is estimated to be more important in the industries that are arguably more cyclical. In other words, the size of macroeconomic effects on default varies across industries in an intuitively reasonable way. For instance, both the output gap and the nominal interest rate are relatively more important in the construction and the real estate sectors in comparison with other industries, and the nominal interest rate is also quite naturally found to be very important for the financial services sector. The macro variables inflation and the real exchange rate are less important from a quantitative perspective, and in most industries coefficients are not statistically significant. However, it is reassuring to find that a depreciating real exchange rate (i.e., lower value, see Figure 2) is associated with a significantly lower default risk in the manufacturing sector, which is the most export-oriented industry. As a robustness check, we examined a model allowing for possibly non-linear relationships between default and the financial ratios and found that the macroeconomic variables are still highly significant and quantitatively important.[18]

Finally, we would like to emphasize that the gain in using firm-specific data for default-risk modelling is substantial. OLS estimates (TSLS give very similar results) for a model of the average quarterly default rate on average financial ratios and the four macro variables are:

$$
\begin{aligned}
df_t = \; & \underset{(0.09)}{-0.15} \; \underset{(0.15)}{+0.15} \left(\frac{\text{EBITDA}}{\text{TA}}\right)_t + \underset{(0.13)}{0.19} \left(\frac{\text{TL}}{\text{TA}}\right)_t + \underset{(0.04)}{0.06} \left(\frac{\text{LA}}{\text{TL}}\right)_t \cdots \\
& \underset{(0.18)}{-0.26} \left(\frac{\text{I}}{\text{TS}}\right)_t \; \underset{(0.04)}{-0.04} \left(\frac{\text{TL}}{\text{TS}}\right)_t + \underset{(0.11)}{0.21} \left(\frac{\text{IP}}{\text{IP+EBITDA}}\right)_t \cdots \\
& \underset{(0.03)}{-0.11} y_{d,t} - \underset{(0.03)}{0.03} \pi_{d,t} + \underset{(0.03)}{0.07} R_{d,t} - \underset{(0.007)}{0.005} q_t + \hat{u}_{df,t},
\end{aligned} \tag{1}
$$

$$R^2 = 0.91, \; \text{DW} = 2.15, \; \text{Sample: } 1990Q1 - 1999Q4 \; (T = 40)$$

If we compare the point estimates for the economy-wide model in Table 3 with those in (1) above, we see that they differ substantially.[19] In particular, the ratios I/TS, LA/TL and TL/TS now enter with counterintuitive signs that have reversed relative to the results in Tables 2 and 3. However, the coefficients for the two key macro variables, the output gap and the nominal interest rate, are very similar to those reported in Table 3 for the economy-wide model. This

[18] When estimating a model where the financial ratios enter in a non-linear way (interaction dummies), we used the cumulated distributions depicted in Figure 1 to categorize the variables (3 categories for each variable). For instance, we classified EBITDA/TA into the decile-based categories $0-10$, $10-90$, $90-100$, whereas TL/TA was classified into the categories $0-75$, $75-90$, $90-100$. This categorization resulted in an increase in pseudo R^2 from 0.35 to 0.42 in the economy-wide model in Table 3. In this model with non-linear balance-sheet variables, the macroeconomic variables still enter highly significantly and with coefficients for the output gap and the nominal interest rate that are very close to those in Table 3. This implies that the macroeconomic variables are still essential for explaining the absolute level of default risk.

[19] The aggregated model in (1) has been estimated without the dummy variables for payment remarks, dividends and failure to submit a financial statement (PAYREMARK, TAXARREARS, PAYDIV, and TTLFS) because they do not enter significantly.

highlights our conclusion that the coefficients for the macroeconomic variables are driven by the time-series dimension of the panel. Since the average financial ratios are quite smooth over time, it is not surprising that we obtain spurious results when the firm-specific information is aggregated. Moreover, some explanatory power is lost by aggregating data; the model in (1) yields an R^2 of 0.91, which can be directly compared with the aggregated fit (see below) of the corresponding model in Table 3, $R^2 = 0.96$.

3.3 Assessing the models' in-sample fit

The last rows in Tables 2 and 3 report on the number of observations, the mean log-likelihood and the pseudo-R^2. The latter measures the ability of the estimated models to explain default at the firm level and is computed using the method of McFadden (1974).[20] Another important and interesting feature of the models is their aggregate performance over time, i.e., how well the models account for the average default frequency. Hence, we report what we label as "aggregate" or "industry" $R^2\,'s$. These are calculated by aggregating all the fitted firm default probabilities in a particular industry model for each quarter $1990Q1-1999Q4$ and then using the resulting 40 time-series observations to compute the implied aggregate R^2.[21] To assess the gain in estimating separate industry-specific models, we also report the pseudo- and industry-R^2 values conditional on the economy-wide model coefficients instead of the industry model coefficients.

By comparing Tables 2 and 3, we see that the pseudo-R^2 is not much affected by the conditioning on macroeconomic factors in any of the industries, merely 1-2 percentage points. However, the industry-R^2 is doubled and sometimes even more than doubled by the introduction of macro variables. Thus, the firm-specific variables account for the cross-section of the default distribution, while the macroeconomic variables in the model play the role of shifting the mean of the default distribution in each period. This also implies that the model with firm-specific information cannot capture the upturns and downturns in the average default rate over time. This is visualized in Figure 3, where we plot the average default rate over time against the fitted values from the economy-wide models in Table 2 (without macro variables) and Table 3 (with macro variables). The results to the right-hand side of the vertical line pertain to out-of-sample results and will be discussed in greater detail in Section 4.1. According to Figure 3, the model

[20] McFadden's (1974) formula for the pseudo-R^2-measure is given by 1-$\frac{\ln L_{\text{model}}}{\ln L_{\text{constant}}}$, where $\ln L_{\text{model}}$ denotes the log-likelihood in the estimated, full model at hand and $\ln L_{\text{constant}}$ is the log-likelihood in an estimated model with only a constant included.

[21] The aggregated R^2 is thus calculated by running a regression of the actual average default rate on the fitted average default rate and a constant.

with both micro and macro variables included appears indeed able to replicate the high default rate during the deep recession/banking crisis in the beginning of the 1990s, as well as the downturn to very moderate default rates during the latter part of the sample. This conclusion is confirmed in Figure 4, where the industry average default rates are plotted together with the average predicted default rates generated by the estimated models in Tables 2 and 3. This finding is very interesting, because it suggests that the extreme default rates recorded during the banking crisis in the early 1990s were not exceptional events that are uninformative in a model context. Rather, they were consequences of unusually bad economic outcomes, both domestically and internationally.[22] An additional feature of interest to note in Tables 2 and 3 is that the fall in pseudo-R^2 values associated with conditioning on the economy-wide model coefficients is distinct but limited, whereas the corresponding reduction in aggregate R^2 is quite substantial. This latter result is confirmed in Figure 5, which shows the average industry default frequencies along with the projected default frequencies using the economy-wide parameter estimates in Table 3. In two cases - the agricultural and the bank, finance & insurance sectors - we note that their industry-R^2 outcomes are negative conditional on the economy-wide model coefficients. At first sight this may seem strange, given that the industry-specific coefficients in Table 3 are not very different from the economy-wide model coefficients. However, as should be clear from Figure 5, these seemingly inconsistent results are driven by the not-reported intercept, which is larger in the economy-wide model compared with the sector models. Therefore it induces a systematically over-prediction of default risk in relation to the actual risk in these sectors.

A conceivable objection to our claim on the importance of conditioning on macroeconomic factors in default risk models is that the significance of these variables simply reflects the fact that the impact of the firm-specific variables changes over time. Accordingly the fit of the models with only firm-specific information would increase dramatically if one were to continuously reestimate the coefficients of the Table 2-models using the most recent quarterly information, thus making the macro variables redundant. Figure 6 displays the estimated coefficients for the financial ratios when allowing for time variation in the economy-wide model.[23] The coefficients

[22] Lindé (2002) shows that a significant share of the variation in domestic macroeconomic variables is of foreign origin.

[23] In these cross-sectional regressions, we impose teh condition that the constant equals the estimated intercept from the economy-wide model in Table 2. The latter model is estimated on the full sample over time and should hence provide a reasonable estimate of the long-run average default level. If the constant is not restricted in the cross-sectional regressions, a substantial amount of the variation in the default rate over time would simply be explained by the intercept. Such variation in the intercept would constitute an improper basis for comparisons with the models including macro variables, since by construction an aggregate R^2 of 100 percent would result. The intercept variations would yield undesirable implications when using the model for scenarioanalysis. Note that we can examine time-varying coefficients only in the economy-wide model, because we do not have a sufficiently large number of defaults in each quarter for some of the industries.We do not report the results of the dummy

for all ratios are highly unstable, switch sign over time - except for the earnings ratio - and do so more pronouncedly toward the latter part of the sample. Moreover, we also find it implausible that earnings were less important during the early 1990s recession, when the default frequency was high. Hence models with only firm-specific variables, which are frequently re-estimated, will inevitably fail to produce adequate out-of-sample forecasts. These are not very appealing features for a default-risk model. Therefore, although a continuously re-estimated model with only firm-specific information will produce a similar-sized aggregate R^2 as a macro-conditional model estimated over time, it does not provide compelling evidence against the claim that macroeconomic variables matter for default risk. To be convincing, a firm-specific model with time-varying coefficients requires an understanding of how the time variation in the coefficients comes about. This appears far-fetched given the economically implausible and irregular patterns displayed in Figure 6. Much sooner one would believe that these patterns point at an omitted-variables problem, namely, the omitted macro variables.

Finally, to further help us understand the role of macro variables for default risk, let us approach the issue from an opposite angle and study the importance of firm-specific variables in the models. One way of demonstrating how much information we lose by omitting the micro data, is to regress the average default frequency on the macroeconomic variables included in Table 3. When doing so we obtain the following result[24]:

$$df_t = \underset{(0.10)}{0.50} - \underset{(0.02)}{0.14} y_{d,t} - \underset{(0.02)}{0.06} \pi_{d,t} + \underset{(0.02)}{0.12} R_{d,t} + \underset{(0.004)}{0.003} q_t + \hat{u}_{df,t},$$

$$R^2 = 0.85,\ \text{DW} = 1.43,\ \text{Sample: } 1990Q1 - 1999Q4\ (T = 40). \tag{2}$$

When comparing this regression with the results in Table 3, we see that we loose about 10 percentage points of the explanatory power in comparison with the economy-wide model when excluding the financial statement variables. Moreover, omitting the firm-specific information introduces mis-specification problems in (2) as indicated by the Durbin-Watson-statistic. This is in contrast to the results in (1), which has a DW-statistic around 2 and hence displays no signs of autocorrelation. The autocorrelation problem in (2) turns out to induce further problems in out-of-sample stability, as documented in Section 4 (see Table 4). Our interpretation is that omitting firm-specific variables when modeling default risk attributes too much of the variation

variables PAYDIV; TTLFS, PAYREMARK and TAXARREARS since they do not exhibit much time-variation.

[24] Here, the dependent variable is defined as a percentage probability of default, i.e., we have multiplied the series in Figure 3 by 100 in order to make the resulting coefficients easier to interpret.

in default risk to the macroeconomic factors in-sample. Since the role of the macroeconomic factors is exaggerated in the estimated model in (2), it will perform less well out-of-sample.

4 Out-of-sample properties of the estimated model

In this section we investigate the robustness of the results in the previous section by examining the out-of-sample properties of the models of Table 2 and Table 3. We use the period $2000Q1 - 2002Q4$, comprising a total of 2,614,248 firm observations in 12 quarters, for the out-of-sample testing. We evaluate the models along two dimensions. First, we study the models' properties at the industry and aggregate level, i.e., we assess their ability to predict future average default rates. The predictions we consider are static one-step-ahead forecasts. For this purpose we have re-estimated all models with the macro variables dated $t-1$, instead of period t as in the case of Table 3, so that we do not have to forecast any of the explanatory variables. Second, we look into the models' properties in predicting future default events at the firm level.

4.1 Evaluating the models at the aggregate and industry level

In Figure 3 and Figures $4-5$, the results to the right-hand side of the vertical bars show the one-step-ahead, out-of-sample performances at the aggregate level and at the industry levels, respectively. As can be seen from Figure 3, the out-of-sample fit in the economy-wide model is remarkably good. Turning to the performance of the industry models in Figure 4, we see that most of the estimated industry models fit the data very well. There are two cases, however, where the models appear to overestimate the default frequency out-of-sample, signalling that the relationship between aggregate fluctuations and the firm default behavior in the sector under consideration has changed. Perhaps not too surprisingly, the two sectors are the ones displaying the highest and most volatile default frequencies during the banking crisis: the hotel and restaurant industry and the real estate industry. So for these sectors, the transmission of aggregate fluctuations into default behavior appears to have changed to some extent.

Before drawing any firm conclusions, it should be noted that these two sectors contain relatively few observations, suggesting that the poor out-of-sample properties could to some degree be attributed to a small-sample problem. This explanation is supported by the fact that the construction sector, which is about four (two) times larger than the hotel and restaurant (real estate) sector and exhibits a strong dependence on aggregate fluctuations according to Table

3, performs well out-of-sample. However, there are other industries that contain even fewer observations than the hotel and restaurant and the real-estate sectors, such as the agricultural and the financial services sectors, but display considerably more stable default frequency patterns throughout the whole sample period and, yet, are not influenced by macro-economic variables to the same extent. The fact that only relatively small sectors perform worse out-of-sample implies that there is no sign of over-prediction when weighting the predicted default probabilities in each industry by their relative size in each period. This becomes clear when we plot the resulting default frequency against the average default frequency in Figure 4 (see lower right diagram). Moreover, by comparing the results in Figure 4 with those in Figure 5, we also see that the out-of-sample fit at the industry level is generally improved by adapting industry-specific models compared to the single aggregate (economy-wide) model. However, at the aggregate level (comparing lower right diagrams), aggregating the weighted results of all the estimated industry models does not seem to offer any gain in comparison to applying the single economy-wide model for all sectors directly. We will be examining this in greater detail below.

In Table 4, we report on the root mean squared prediction errors (RMSEs), one-step-ahead, for the estimated models of Tables 2 and 3. We also show results for three reference time series models: a random walk model, a 4-quarter moving-average model, and the model estimated on aggregate data with only macroeconomic data included (eq. 2, denoted "Industry OLS macroregression").[25] The results in Table 4 pertain to default risk models that have been re-estimated using macro variables that are lagged one quarter. This ensures that all models in Table 4 have been estimated on the same information, thereby allowing for a fair comparison between the logit and the time series models. In the "Industry OLS macroregression" models an additional dummy for the third quarter is included. Finally, it is imperative to notice that the RMSEs are shown in percent, i.e., the actual and fitted default frequencies have been multiplied with a factor of 100 before the prediction errors are calculated.

From inspection of Table 4, it is evident from the first row in the lower panel that the effect on forecasting performance from conditioning on both macro and firm-specific information is considerable. The largest gain is found for the economy-wide model where forecast precision increases by as much as a factor of seven when we include macro variables. The corresponding factors for the industry-specific models range between 2 and 5, disregarding the not classified

[25] We also experimented with estimated AR(p), $p = 1, ...4$, models. These were found to be inferior to the models reported here, presumably due to the downward shift in default frequencies between the in-sample and out-of-sample periods.

residual industry, where there is barely any improvement. This can be interpreted as evidence of an effect from macro variables over and above some spurious industry effect, i.e., controlling for industry belonging will not shut down the influence from macro variables. Moreover, the industry-specific models generate lower RMSEs compared with the industry models conditional on coefficients from the economy-wide model in Table 3, except for the retail sector. In the retail sector the industry-specific model has an RMSE of 0.12 percent and the industry model based on economy-wide coefficients has a much smaller RMSE of 0.05.

By and large, the above findings constitute evidence that the industry-specific models are not over-parameterized with respect to macroeconomic variables. Therefore it will typically be worthwhile to work with an industry-specific model if the focus is on understanding default behavior in a particular industry. However, if the interest is modeling aggregate default behavior only, the economy-wide default model appears to suffice. This tentative conclusion can be drawn from the two right-most columns of the second row in Table 4. There the different industry forecasts computed with the industry-specific models have been weighted to a forecast for the aggregate default frequency. This results in a slightly higher RMSE in comparison with the RMSE for the economy-wide model (0.066 and 0.0478, respectively). Although this difference in RMSE is very low in comparison with the other models in absolute terms, it is rather high in relative levels. This in its turn becomes clear from inspection of the RMSE ratios presented in the second row of the lower panel (0.724). This begs the question as to why the industry models are inferior to the economy-wide model in terms of out-of-sample forecasting performance for the whole economy when they outperform the economy-wide model in almost every industry. The difference can be shown to be driven by the relatively large retail industry, which suffers from out-of-sample over-prediction for the industry-specific retail model. For the economy-wide model (compare the graphs for Retail in Figures 4 and 5) such an over-prediction does not occur.[26]

Comparing the industry models in Table 3 with the time series models, we also see that while the random walk model is doing better in four out of ten sectors, and the four-quarter moving-average specification is better five out of ten times at the industry level, they are still both clearly inferior to the aggregated industry models. This implies that they are also inferior

[26] As a check of the validity of this claim we re-calculated the RMSE for the aggregated industry and economy-wide models excluding the retail industry altogether. The results are 0.0469 for the industry aggregate model and 0.0525 for the economy-wide model. This confirms that the retail sector has a positive influence on the RMSE for the economy-wide model and likewise a negative influence on the industry aggregate RMSE for the out-of-sample period we consider here.

in terms of RMSE fit to the aggregate model specification in Table 3 (which conditions on aggregate fluctuations). The models that are based on OLS regressions for average industry default frequencies on the macro variables only (Industry OLS macro regressions, see eq. 2) also perform poorly out-of-sample in comparison with the Table 3 models. As already discussed in Section 3.3, the unfavorable performance out-of-sample for the models estimated on aggregated variables only is most likely driven by the tendency for such models to erroneously attribute too much of the fluctuations in the default frequencies in-sample to fluctuations in the macro variables; this is a consequence of the firm-specific variables being incorrectly omitted.

To sum up, we have found strong evidence that the favorable fit in-sample of the estimated industry (and aggregate) models, conditional on macro variables, is preserved out-of-sample at the industry and aggregate level. This suggests that the macroeconomic factors that enter into the model are structural, and not merely improving the in-sample fit of the models. An important reason why the favorable out-of-sample performance is reassuring for the hypothesis that aggregate variables matter is that the in-sample and the out-of sample periods taken together cover several upturns and downturns in the Swedish economy. This is evident from the output-gap series in Figure 2. Finally, we have also documented that there are only small gains in terms of forecasting accuracy to be made by using industry-specific models rather than simply an aggregate model, as long as an appropriate set of macroeconomic variables is included.

4.2 Evaluating the models at the firm and the industry level

In this subsection, we turn to the out-of-sample properties of the estimated models at the micro level, i.e., their ability to predict default events at the firm level. In particular, we evaluate the models' performance in terms of ranking relative firm riskiness, as well as their determination of firms' absolute risk. In addition, for the out-of-sample period, we report the industry-specific pseudo-R^2 conditional on the industry-specific model coefficients of Table 3, as well as the pseudo-R^2 calculated conditional on the economy-wide model coefficients. The results are displayed in Table 5.

First, starting with the pseudo-R^2 for the models with industry-specific coefficients and comparing the in-sample and out-of-sample results reported in Tables 3 and 5, respectively, we see that the explanatory power out-of-sample is in fact either higher than in-sample or unchanged in five out of ten industries. The lower panel of Table 5 shows, for the economy-wide model, that the explanatory power has increased substantially from 0.35 to 0.39. In the three cases where the pseudo-R^2 decreases (hotel and restaurant, transportation and real-estate), it does so only marginally.

Next we turn to the pseudo-R^2 for the predictions based on the economy-wide model coefficients. In the lower panel of Table 5, we see, relative to Table 3, that the explanatory power has increased in all industries except for agriculture, where it is unchanged. However, the pseudo-R^2 values generated when using industry-specific coefficients are typically at least as large as the ones obtained when using economy-wide model coefficients, with the exception of the real-estate sector (compare upper and lower panels of Table 5). This implies that pseudo-R^2 at the aggregate level is slightly lower for the economy-wide model compared with an aggregation of pseudo-R^2 over the industry-specific models (denoted Industry aggregate in Table 5). These results provide support for two important conclusions. First, the industry models are not over-parameterized. Second, the reduced-form coefficients appear to be stable over time and the regressions thus reflect steady relationships that hold even out-of-sample.

Third, moving on to measures of relative risk, we follow Shumway (2001). Another important scale along which to evaluate the models is their ability to rank firms according to their riskiness in terms of *ex post* default frequencies. In other words, we investigate if the estimated default risk models assign the largest *ex ante* default probabilities to the riskiest firms, and vice versa for the least risky firms. At a first glance, we see from Table 5 that the estimated models classify roughly $75 - 80$ percent of the defaulting firms in the first decile. These numbers are about the same as those reported in-sample by Shumway for a data set that was substantially smaller and included only listed firms. Our models cover the entire population of Swedish incorporated businesses, of which only a very small subset is listed on the stock exchange (about 500 out of $250,000$). We therefore conclude that our models are quite successful in ranking firms according to their level of default risk. The empirical performance of the models constitutes an important support for our conclusion that the role of macroeconomic variables in models of default risk is not driven by the omission of key microeconomic variables.

Table 5 also reveals that the quality of the risk rankings does not depend on whether we condition on industry-specific coefficients or coefficients from the economy-wide model. This contrasts with our findings in the previous subsection, where we found that conditioning on industry-specific parameters improved the models' empirical performance at the industry level. The explanation for these seemingly inconsistent results lies in the fact that the most important difference between the economy-wide and industry-specific models is made up by the varying impact of the aggregate factors. Those factors have little impact on the firms' riskiness ranking and hence their inclusion or omission has little impact on the models' ability to risk-rank firms.[27]

Finally, to assess the out-of-sample properties of the models at the microeconomic level in a more formal and absolute sense, we compare the distribution of estimated out-of-sample default probabilities with the actual default frequencies. We do so by sorting all estimated default probabilities according to increasing size and calculating the average probability of default in each percentile. We then compare these percentiles with the actual default experience of these percentiles.[28] In Figure 7, we plot the result where we have used both the industry-specific and the economy-wide model coefficients in Table 3 to compute the estimated default probabilities for each firm. On the x-axis, we have the estimated default frequency in a given percentile, and on the y-axis, we have the actual default frequency in each percentile. In the figure each dot is a percentile, and in order to make the results easier to access, a logarithmic scale is used for both the estimated and actual series. If the estimated models could perfectly predict the absolute riskiness of the firms within each percentile, all dots would line up along the 45-degree line drawn in the figures, which has a slope of unity and intercept equal to nil. As can be seen in Figure 7, this is not the case for either model, but the dots are generally very close to the line, suggesting that the absolute riskiness ranking is very accurate. In particular, the models that include macroeconomic factors appear to better capture the absolute risk level, since the models without macroeconomic variables tend to overestimate default risk. The mean standard deviations from the 45-degree line (in logarithmic scale) are 0.82 and 0.90 for the upper and lower left panels (no macro variables) and 0.63 for both the upper and lower right panels (with macro variables). However, since the standard errors for the industry aggregate and economy-wide model (that both condition on macroeconomic factors) are about the same, the results suggest that aggregation of the industry-specific models does not outperform the economy-wide model in this dimension. This confirms our previous findings: it is sufficient to condition on aggregate factors in the economy-wide model to obtain an acceptable model at the aggregate level, but industry-specific models are typically superior for predicting default risk at the industry level.

[27] In a given quarter, aggregate shocks have zero influence on the ranking because they affect the default probabilities equally much by the way the estimated models are constructed.

[28] It would have been very interesting to report results for the different industries as well, but there are not enough defaults out-of-sample to split up the data in percentiles for each industry.

5 Conclusions

In this paper, we study the interaction between macroeconomic fluctuations and default risk at the firm level using reduced-form methods. To this end we collected a large panel data set for the Swedish economy during $1990 - 2002$, a period covering a deep recession, and an associated banking crisis in the early 1990s, followed by a boom in the latter part of the 1990s, as well as a downturn in the beginning of the 2000s. We divided the sample in two parts, $1990 - 1999$ and $2000 - 2002$. We use the former period for model estimation and the latter to provide an assessment of whether the impact of aggregate fluctuations on default risk, over and above firm-specific variables, is a robust regularity valid for the entire sample period.

We present four main findings. First, we provide insight into the significance of aggregate fluctuations for defaults among not only listed but even privately held firms. This is of significance, since privately held businesses typically account for over half of GDP in developed economies. Second, a nearly exhaustive set of firm-specific background variables permits us to investigate the importance of and interaction between firm-specific variables and macroeconomic information - a nearly unexplored area. Third, we document that a simple logit approach to model default at the firm level, using both firm-specific and macroeconomic variables, can explain the peaking default frequencies during the Swedish banking crisis of the early 1990s as well as well for the considerably lower default frequencies in the late 1990s. The length of our panel enables us to do extensive out-of-sample performance tests of our model. The estimated models are shown to be very robust and successful out-of-sample, suggesting that aggregate fluctuations play a truly prominent role in understanding the absolute level of firm default risk. Finally, the width of our panel permits us to investigate the relation between aggregate fluctuations and firm defaults across industries. This shows that macroeconomic variables have a robust and "structural" impact on business defaults.

We do not interpret our results as implications of aggregate fluctuations being the most important source of default at the firm level. Rather, we argue that the results suggest that macroeconomic factors shift the mean of the default risk distribution over time and thereby are the most important source of the level of default risk.

In view of these results, we conclude by providing some suggestions as to why aggregate fluctuations should be expected to have a statistically important impact on firm default behavior, over and above the effect that firm-specific variables, which themselves move in response to macroeconomic fluctuations, have. Hackbarth, Miao and Morellec (2006) argue that the dependence of cash flows on economic conditions lead to firms' optimal default thresholds being affected by aggregate shocks. Hence aggregate fluctuations can trigger simultaneous defaults. Another argument for why aggregate variables might contain predictive information for firm-default risk over and above the firm-specific information is related to the costliness of monitoring. If monitoring borrowers is costly for banks, then banks may use aggregate information to assess the probability of getting repayment on loans granted. That is, banks may form their credit-granting policies on the basis of macroeconomic forecasts and decide to not extend new lines of credit to firms with a given set of performance indicators in one particular phase of the business cycle, but readily do so in another phase. In other words, banks resort to using the macroeconomic stance in their decision processes. Yet another argument follows a similar line of reasoning. If entrepreneurs have imperfect information about their own future business prospects, they may resort to using aggregate conditions as a basis for their decision to either invest more effort in a firm or declare bankruptcy. In addition, if firms are borrowing-constrained, then the nominal interest rate will be an important and direct determinant of default risk. A final possibility is that firms may be inclined to adjust their yearly accounts to, e.g., smooth profit over time in order to please banks' monitoring efforts, and thereby reduce the predictive power of firm-level information. We believe that formalizing the theory of how macroeconomic variables affect firm defaults and assessing the empirical plausibility of the arguments above are important issues for future research.

References

Altman, Edward I., (1968), Financial ratios, discriminant analysis and the prediction of corporate bankruptcy, *Journal of Finance*, Vol. 23 No. 4, pp. 589-611.

Altman, Edward I., (1971), Railroad bankruptcy propensity, *Journal of Finance*, Vol. 26 No. 2, pp. 333-345.

Altman, Edward I., (1973), Predicting railroad bankruptcies in America, *Bell Journal of Economics*, No. 4 (1), pp. 184-211.

Altman, Edward I., (1984), The success of business failure prediction models, *Journal of Banking and Finance*, No. 4, pp. 171-198.

Altman, Edward I., and Anthony Saunders, (1997), Credit risk measurement: developments over the last twenty years, *Journal of Banking and Finance*, No. 21 (11-12), pp. 1721-42.

Altman, Edward I., and Anthony Saunders, (2001), An analysis and critique of the BIS proposal on capital adequacy and ratings, *Journal of Banking and Finance*, No. 25 (1), pp. 25-46.

Altman, Edward, Frydman, Halina Frydman and Duen-Li Kao, (1985), Introducing recursive partitioning for financial classification: the case of financial distress, *Journal of Finance*, Vol. XL, No. 1, pp. 269-291.

Arellano, Manuell, and Stephen Bond, (1991), Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations, *Review of Economic Studies*, Vol. 58, No. 2, pp. 277-297.

Ayyagari, Meghana, Thorsten Beck and Asli Demirguc-Kunt, (2007), Small and medium enterprises across the globe: a new database, *Small Business Economics* 29, pp. 415-34..

Bernanke, Ben, (1983), Nonmonetary effects of the financial crisis in the propagation mechanism of the Great Depression, *American Economic Review* Vol. 73 No.3, pp. *257-276.*

Bernanke, Ben, (1993), Nonmonetary effects of the financial crisis in the propagation mechanism of the Great Depression, *Federal Reserve Bank of New York Quarterly Review* 18(1), pp. *50-70.*

Bernanke, Ben, and Mark Gertler, (1989), Agency costs, net worth and business fluctuations, *American Economic Review* Vol. 79 No.1, pp. *14-31.*

Bernanke, Ben, and Mark Gertler, (1990), Financial fragility and economic performance, *Quarterly Journal of Economics* 105(1), pp. 87-1*14.*

Bernanke, Ben, and Mark Gertler, (1995), Inside the black box: the credit channel of monetary policy transmission, *Journal of Economic Perspectives 9*(4), pp. 27-*48.*

Bernanke, Ben, Mark Gertler and Simon Gilchrist, (1996), The financial accelarator and the flight to quality, *Review of Economics and Statistics* Vol. 78 No.1, pp. *1-15.*

Bharath, Sreedhar T., and Tyler Shumway, (2008), Forecasting default with the Merton distance to default model, *Review of Financial Studies* 21 (3), pp. 1339-1369.

Blume, Marshall E., Donald B. Keim, and Sandeep A. Patel, (1991), Returns and volatility of low-grade bonds 1977-1989, *Journal of Finance* Vol. 46 No. 1, 49–74.

Calomiris, Charles, and R. Glenn Hubbard, (1989), Price flexibility, credit availability and economic fluctuations: evidence from the U.S., 1894-1909, *Quarterly Journal of Economics* 104(3), pp. 429-*452.*

Carey, Mark, (1998), Credit risk in private debt portfolios , *Journal of Finance,* Vol. 53 No. 4, pp. 1363-1387.

Carey, Mark and Mark Hrycay, (2001), Parameterizing credit risk models with rating data, *Journal of Banking and Finance,* 25, pp. 197-270.

Carling, Kenneth, Tor Jacobson, Jesper Lindé and Kasper Roszbach, (2007), Corporate credit risk modelling and the macroeconomy, *Journal of Banking and Finance* Vol. 31, No. 3, pp. 845-868.

Christiano, Lawrence J., Eichenbaum, Martin and Charles L. Evans, (1999), Monetary policy shocks: what have we learned and to what end?, Chapter 2 in John B. Taylor and Michael Woodford (Eds.), *Handbook of Macroeconomics*, Vol. 1A, Elsevier Science, Amsterdam.

Christoffersen, Peter F., (1998), "Evalutating interval forecasts", *International Economic Review*, Vol. 39, No. 4, pp. 841-862.

Christiano, Lawrence J., Eichenbaum, Martin and Charles L. Evans, (2005), "Nominal rigidities and the dynamic effects of a shock to monetary policy", *Journal of Political Economy*, forthcoming.

Coe, Patrick, (2002), Financial crisis and the Great Depression: a regime switching approach, *Journal of Money, Credit and Banking* 34(1), pp. 76-93.

Duffie, Darell, Leandro Saita and Ke Wang, (2007), Multi-period corporate default prediction, *Journal of Financial Economics* 83 (3), pp. 635-665.

Englund, Peter, (1999), The Swedish banking crisis: roots and consequences, *Oxford Review of Economic Policy*, 15 (3) Autumn , pp. 80-97.

Fons, Jerome, (1991), *An Approach to Forecasting Default Rates*, Working Paper, Moody's Investors Service.

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.

Gertler, Mark, and Simon Gilchrist, (1994), Monetary policy, business cycles and the behavior of small manufacturing firms, *Quarterly Journal of Economics* 109, pp. 309-340.

Gordy Michael B., (2000a), A comparative anatomy of credit risk models, *Journal of Banking and Finance*, 24, No 1-2, pp. 119-149.

Hackbarth, Dirk, Jianjun Miao and Erwan Morellec, (2006), Capital structure credit risk and macroeconomic conditions, *Journal of Financial Economics* 82 (3), pp. 519-550

Helwege, J. and P. Kleiman, (1997), Understanding aggregate default rates of high-yield bonds, *Journal of Fixed Income Vol.* 7, pp. 55–62.

Jacobson, Tor and Jesper Lindé, (2000), Credit rating and the business cycle: can bankruptcies be forecast?, *Sveriges Riksbank Economic Review*, 2000, No 4, pp. 11-33.

Jacobson, Tor, Jesper Lindé, and Kasper Roszbach, (2005), Exploring interactions between real activity and the financial stance, *Journal of Financial Stability* 1 (3), pp. 308-341.

Jonsson, J. and M. Fridson, (1996), Forecasting default rates on high-yield bonds, *Journal of Fixed Income*, pp. 69–77.

Kobe, Kathryn, (2007), The small business share of GDP 1998-2004, Manuscript No. 299, Small Business Administration, http://www.sba.gov/ADVO/research/rs299.pdf.

Laitila, Thomas, (1993), A Pseudo-R^2 measure for limited and qualitative dependent variable models, *Journal of Econometrics*, No. 56, pp. 341-356.

Lindé, Jesper, (2002), Monetary policy shocks and business cycle fluctuations in a small open economy: Sweden 1986-2002, Working Paper No. 153, Sveriges Riksbank, Stockholm.

Lowe, Philip, (2001), Maintaining financial stability: possible policy options, *Sveriges Riksbank Economic Review* 2001(2), pp. *25-33.*

Ludvigson, Sidney, (1998), The channel of monetary transmission to demand: evidence from the market for automobile credit, *Journal of Money, Credit and Banking*, Vol. 30 No. 3, pp. 365-383.

McFadden, Daniel, (1974), The measurement of urban travel demand, *Journal of Public Economics* 3(4), pp. 303-328.

Merton, Robert C., (1974), On the pricing of corporate debt: the risk structure of interest rates, *Journal of Finance* Vol. 29, pp. 449-470.

Peek, Joe, and Eric Rosengren, (2000), Collateral damage: effects of the Japanese bank crisis on real activity in the United States, *American Economic Review* 90(1), pp. 30-45.

Pesaran, M. Hashem, Til Schuermann, Björn-Jakob Treutler, and Scott M. Weiner, (2006), Macroeconomic dynamics and credit risk: a global perspective, *Journal of Money, Credit and Banking*, 38 (5), pp. 1211-1261.

Platt, Harlan D., and Marjorie B. Platt, (1991), A note on the use of industry-relative ratios in bankruptcy prediction, *Journal of Banking and Finance*, Vol 15, pp. 1183-1194.

Poole, William, (1993), Credit veils and credit realities, *Federal Reserve Bank of New York Quarterly Review* 18(1), pp. 83-85.

Repullo, Rafael, and Javier Suarez, (2000), Entrepreneurial moral hazard and bank monitoring: a model of the credit channel, *European Economic Review* 44, pp. 1931-1950.

Samolyk, Katherine, (1994), Banking conditions and regional economic performance: evidence of a regional credit channel, *Journal of Monetary Economics* 34(2), pp. 259-278.

Shumway, Tyler, (2001), Forecasting bankruptcy more accurately: a simple hazard model, *Journal of Business*, Vol. 74, No 1, pp. 101-124.

Wilson, Thomas, (1997), Portfolio credit risk (I), *Risk*, Vol. 10, No. 9.

6 Data appendix

As mentioned in Section 2.1, the default definition we adopt is the following: a firm is considered to have a default status once any of the following events occurs: the firm is declared legally bankrupt, has suspended payments, has negotiated a debt composition settlement, is undergoing a re-construction, or is distraint without assets. The data we use to construct the default variable have been provided by Upplysningscentralen AB (UC), the main Swedish credit bureau, that is jointly owned by (most of) the Swedish banks. UC taps its information from *Tingsrätten*, District Court, *Bolagsverket*, the Swedish Companies Registration Office (SCRO), and *Kronofogdemyndigheten*, the Swedish Enforcement Authority.

UC stores information on minor and major distress events in two different databases, AM and JP. In the first database, variable names are constructed by giving each event the name AMTYP, a Swedish acronym for remark type, and an integer number suffix. For example, AMTYP12 is a dummy variable that indicates if a firm has suspended its payments. The variables we use to construct our default variable are "declared bankrupt," "declared minor bankruptcy," "suspension of payments," "debt composition settlement decided," "company reconstruction started," "company reconstruction concluded," and "distraint without assets." The second database contains further information, 27 variables in all, on various milestones and stages in a broader category of major (mostly but not exclusively distress) events for each registered firm. The foremost variables we use from this database are related to legal bankruptcy: "bankruptcy procedures started," "bankruptcy procedures concluded," "bankruptcy procedures concluded with a surplus," "bankruptcy procedures continued," and "declared bankrupt." In addition we use: "negotiations on a debt composition settlement started," and "negotiations on a debt composition settlement concluded."

If any of the above distress-event dummy variables equals one at some moment in our sample period, the firm in question is considered to have defaulted in that particular quarter. In the following quarter, we let the firm exit our data set. If more than one of these distress events are observed for a specific firm over our sample period, we assume the firm in question has defaulted in the quarter during which the first of these events took place. An additional variable we use from the second data set indicates if a "bankruptcy [was] cancelled" by a court. Over the whole sample period (i.e., in-sample and out-of-sample) this occurs 11 times, and seven of these 11 events relate to firms that default later on. We treat firms for which the bankruptcy status was cancelled by the District Court as healthy until the data indicate otherwise. Moreover, we let firms that default but re-emerge from their default status exit the data set after the quarter in which default takes place; they re-enter in the quarter in which UC registered that the default status had been "removed."

Our decision to let firms that default exit the data set in the subsequent quarter is based on the following statistics: of all 123,023 defaults in our whole data set 117,481 are terminal in the sense that no new information on the firms is added to any of the databases.[29] The remaining observations concern firms that default twice within the sample period. Of these observations, 3,555 defaults are terminal at the second occurrence, while 107 re-emerge even after the second default. No firm defaults more than two times within our sample period.

Of the 117,481 first-time-is-terminal defaults 111,702 are legal bankruptcy declarations. For about 45 percent of these firms another default-triggering distress event occurs simultaneously, i.e. during the same quarter, in our data. In most cases this is the variable "bankruptcy proceedings started." Nearly all of the remaining terminal defaults, i.e., those that are not bankruptcies, are associated with "distraint, no assets." The remaining distress events account for less than 1 percent of the first-time-is-terminal defaults.

For the firms that re-emerge after a default, the first default involves a legal bankruptcy in less than half a percent of all cases and "distraint, no assets" in 98 percent. At their second default, these percentages are reversed for the terminal defaults. Among the firms that experience a second non-terminal default, 98 percent cause their second default by obtaining the "distraint, no assets" status.

[29] Firms that are declared bankrupt at some point do not disappear from the databases that UC maintains. Company numbers (*organisationsnummer*) are unique and never re-used by the tax authorities.

Table 1: Descriptive statistics for firm-specific micro data 1990Q1-1999Q4
Panel A: Non-truncated data

	Agriculture	Manufacturing	Construction	Retail	Hotel & Resturant	Transport	Bank, Finance & Insurance	Real-Estate	Consulting & Rental	Not Classified	Total
Defaulted											
Number Obs	1455	11730	10971	30896	5302	5412	708	8650	15353	15092	105569
EBITDA/TA	-0.52	-3.93	-0.46	-9.24	-1.60	-1.13	-400	-1.72	-4.97	21.2	-5.84
	(6.67)	(241)	(28.4)	(807)	(74.9)	(46.5)	(9270)	(550)	(449)	(2380)	(1190)
TL/TA	4.55	307	3.54	148	3.70	3.66	419	25.1	18.0	1314	201
	(62.7)	(27900)	(100)	(19000)	(65,1)	(44,1)	(9260)	(1160)	(925)	(73400)	(25500)
LA/TL	0.39	0.66	0.25	0.41	0.27	0.47	0.44	0.26	0.69	1.90	0.57
	(4.88)	(35.9)	(3.57)	(23.4)	(7.42)	(10.1)	(2.28)	(3.77)	(16.0)	(49.8)	(24.1)
I/TS	0.60	198	0.45	3.31	0.05	0.31	2.60	3.99	-0.98	1.94	26.2
	(2.86)	(17700)	(9.84)	(350)	(0.27)	(9.68)	(52.5)	(219)	(207)	(42.1)	(6250)
TL/TS	1.05	0.24	1.31	0.29	0.57	0.38	2.12	2.94	2.36	1.77	0.96
	(4.10)	(0.82)	(9.78)	(3.74)	(2.20)	(3.75)	(15.1)	(15.4)	(32.9)	(16.7)	(14.90)
IP/(IP+EBITDA)	-0.22	0.41	0.07	0.30	-0.14	1.79	0.09	0.11	0.42	0.03	0.32
	(17.5)	(24.0)	(13.9)	(15.8)	(14.61)	(95.0)	(8.16)	(18.2)	(23.0)	(7.75)	(28.34)
PAYDIV	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.00	0.01
	(0.07)	(0.08)	(0.09)	(0.08)	(0.06)	(0.08)	(0.14)	(0.08)	(0.10)	(0.06)	(0.08)
TTLFS	0.41	0.30	0.37	0.36	0.35	0.41	0.52	0.40	0.44	0.51	0.39
	(0.49)	0.46	(0.48)	(0.49)	(0.48)	(0.49)	(0.50)	(0.49)	(0.50)	(0.50)	(0.49)
PAYREMARK	0.15	0.11	0.15	0.14	0.17	0.17	0.22	0.14	0.17	0.18	0.15
	(0.36)	(0.32)	(0.36)	(0.35)	(0.37)	(0.38)	(0.41)	(0.35)	(0.38)	(0.38)	(0.36)
TAXARREARS	0.45	0.35	0.46	0.39	0.46	0.50	0.47	0.36	0.45	0.39	0.41
	(0.50)	(0.48)	(0.50)	(0.49)	(0.50)	(0.50)	(0.50)	(0.48)	(0.50)	(0.49)	(0.49)
Non-defaulted											
Number Obs	218149	1023629	837687	2130768	246391	510262	98655	465181	1724537	745310	8000569
EBITDA/TA	0.19	0.17	0.09	-0.21	-0.19	-0.28	0.01	0.66	0.05	-1.56	-0.14
	(9.18)	(88.5)	(23.1)	(169)	(27.7)	(206)	(17.6)	(200)	(59.9)	(608)	(220)
TL/TA	1.78	1.37	0.88	2.11	1.47	1.71	1.46	13.3	2.42	9.77	3.25
	(230)	(159)	(33.6)	(313)	(46.1)	(440)	(36.3)	(2647)	(714)	(3470)	(1290)
LA/TL	0.65	0.71	0.49	0.56	1.60	1.13	2.40	1.21	1.64	2.57	1.12
	(8.95)	(22.0)	(16.8)	(105)	(223)	(142)	(38.4)	(87.3)	(122)	(185)	(110)
I/TS	0.45	0.84	1.42	6.52	0.13	0.12	1.80	39.0	0.68	1.50	4.58
	(12.1)	(160)	(400)	(2680)	(16.9)	(30.0)	(327)	(2460)	(270)	(218)	(6070)
TL/TS	1.10	1.12	1.11	0.54	0.67	1.30	24.4	7.76	3.72	1.80	1.97
	(13.7)	(18.1)	(21.2)	(12.9)	(10.4)	(35.7)	(425)	(75.6)	(97.0)	(49.0)	(70.3)
IP/(IP+EBITDA)	-3.0	-4.62E+10	2.35E+09	-8.59E+10	0.21	-1.16E+10	0.15	0.27	1.30E+10	1.37E+11	-1.40E+10
	(738)	(2.27E+13)	(2.14E+12)	(4.62E+13)	(30.7)	(8.24E+12)	(12.7)	(18.5)	(2.52E+13)	(6.65E+13)	(3.44E+13)
PAYDIV	0.16	0.15	0.13	0.13	0.06	0.13	0.14	0.10	0.16	0.11	0.14
	(0.36)	(0.35)	(0.33)	(0.33)	(0.23)	(0.33)	(0.34)	(0.29)	(0.37)	(0.31)	(0.34)
TTLFS	0.01	0.01	0.01	0.01	0.02	0.01	0.02	0.02	0.02	0.03	0.01
	(0.09)	(0.10)	(0.09)	(0.11)	(0.13)	(0.11)	(0.14)	(0.12)	0.12	(0.17)	(0.12)
PAYREMARK	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00
	(0.05)	(0.06)	(0.06)	(0.06)	(0.08)	(0.06)	(0.06)	(0.07)	(0.05)	(0.05)	(0.06)
TAXARREARS	0.03	0.04	0.04	0.03	0.06	0.04	0.02	0.03	0.03	0.03	0.03
	(0.16)	(0.19)	(0.20)	(0.18)	(0.24)	(0.19)	(0.14)	(0.17)	(0.17)	(0.16)	(0.18)

Notes: The definition of variables are: EBITDA = earnings before taxes, interest payments and depreciations; TA = total assets; TL = total liabilities; LA = liquid assets; I = inventories; TS = total sales; IP = sum of net interest payments on debt and extra-ordinary net income; PAYDIV = a dummy variable equal 1 if the firm has paid out dividends during the accounting period and 0 otherwise; TTLFS = a dummy variable equal to 1 if the firm has not submitted an annual report in the previous year, and 0 otherwise; PAYREMARK = 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; (i) a "non-performing loan" at a bank, or (ii) a bankruptcy petition, or (iii) issuance of a court order to pay a debt, or (iv) seizure of property; TAXARREARS = a dummy variable taking the value of 1 if the firm is in various tax arrears.

Table 2: Regression results 1990Q1-1999Q4 for the default risk model estimated with only firm-specific variables

	Agriculture	Manu-facturing	Construction	Retail	Hotel & Restaurant	Transport	Bank. Finance & Insurance	Real Estate	Consulting & Rental	Not Classified	Economy Wide
Firm-specific variables [a]											
EBITDA/TA	-1.308	-1.419	-1.472	-0.957	-0.856	-1.148	-0.361	-0.738	-0.857	-1.069	-0.949
	(0.115)	(0.045)	(0.053)	(0.024)	(0.040)	(0.056)	(0.098)	(0.059)	(0.030)	(0.028)	(0.012)
TL/TA	0.989	1.104	0.599	0.636	0.205	0.753	0.185	0.726	0.342	0.160	0.491
	(0.082)	(0.034)	(0.041)	(0.016)	(0.028)	(0.046)	(0.054)	(0.030)	(0.023)	(0.021)	(0.008)
LA/TL	-0.317	-0.488	-0.493	-0.373	-0.092	-0.192	-0.180	-0.317	-0.247	0.011	-0.251
	(0.093)	(0.040)	(0.042)	(0.020)	(0.041)	(0.036)	(0.052)	(0.035)	(0.017)	(0.009)	(0.008)
I/TS	0.069	0.325	-0.177	0.274	1.315	0.040	0.014	0.053	0.340	0.083	0.124
	(0.049)	(0.036)	(0.044)	(0.016)	(0.310)	(0.240)	(0.055)	(0.009)	(0.041)	(0.021)	(0.006)
TL/TS	0.177	0.128	0.306	0.157	0.237	0.091	0.038	0.068	0.202	0.358	0.164
	(0.025)	(0.006)	(0.008)	(0.004)	(0.013)	(0.010)	(0.015)	(0.006)	(0.006)	(0.006)	(0.002)
IP/(IP+EBITDA)	0.094	0.103	0.055	0.061	0.003	0.194	0.070	0.180	0.045	0.145	0.088
	(0.037)	(0.013)	(0.014)	(0.007)	(0.019)	(0.025)	(0.052)	(0.017)	(0.012)	(0.014)	(0.004)
PAYREMARK	1.284	1.449	1.691	1.523	1.531	1.682	2.239	1.604	1.775	2.512	1.712
	(0.123)	(0.045)	(0.045)	(0.028)	(0.058)	(0.061)	(0.157)	(0.053)	(0.036)	(0.044)	(0.015)
TAXARREARS	2.796	2.216	2.461	2.449	2.380	2.837	3.108	2.419	2.848	2.693	2.566
	(0.078)	(0.028)	(0.029)	(0.017)	(0.040)	(0.041)	(0.110)	(0.033)	(0.024)	(0.027)	(0.009)
PAYDIV	-2.310	-1.912	-1.900	-2.119	-1.667	-1.755	-1.180	-1.710	-1.754	-2.204	-2.004
	(0.401)	(0.114)	(0.111)	(0.076)	(0.219)	(0.178)	(0.296)	(0.139)	(0.085)	(0.132)	(0.039)
TTLFS	4.161	3.695	4.046	3.643	3.371	3.918	3.720	3.615	3.796	3.333	3.670
	(0.072)	(0.027)	(0.029)	(0.016)	(0.039)	(0.038)	(0.097)	(0.030)	(0.022)	(0.022)	(0.009)
Mean log-likelihood	-0.024	-0.044	-0.044	-0.051	-0.071	-0.036	-0.025	-0.061	-0.032	-0.058	-0.046
Pseudo R^2	0.39	0.29	0.36	0.32	0.31	0.38	0.40	0.33	0.38	0.41	0.34
Pseudo R^2 | agg.par. [b]	0.35	0.28	0.36	0.31	0.30	0.38	0.32	0.32	0.37	0.39	0.34
Industry R^2	0.50	0.51	0.47	0.53	0.52	0.38	0.27	0.52	0.45	0.56	0.49
Industry R^2 | agg.par. [b]	-2.09	0.39	0.37	0.42	0.39	0.27	-1.97	0.49	0.02	0.46	0.49
Number of obs	219,604	1,035,359	848,658	2,161,664	251,693	515,674	99,363	473,831	1,739,890	760,402	8,106,138

Notes: Standard errors in parentheses. The variables are not scaled, so the importance of a variable cannot be interpreted directly from the size of the parameter estimate. [a] See Subsection 2.1 for exact definition of these variables. [b] Pseudo R^2. | agg.par is the Pseudo R^2 value calculated for each industry using the estimated coefficients in the economy-wide model (i.e., the coefficients in the last column in the table above). The pseudo R^2 values are calculated according to McFadden (1974).

Table 3: Regression results 1990Q1-1999Q4 for the default risk model estimated with both firm-specific and aggregate variables

	Agriculture	Manu-facturing	Construction	Retail	Hotel & Restaurant	Transport	Bank, Finance & Insurance	Real Estate	Consulting & Rental	Not Classified	Economy Wide
Firm-specific variables [a]											
EBITDA/TA	-1.323	-1.412	-1.420	-0.950	-0.850	-1.159	-0.373	-0.673	-0.880	-1.073	-0.954
	(-0.025)	(0.006)	(0.008)	(0.004)	(0.013)	(0.011)	(0.015)	(0.006)	(0.006)	(0.006)	(0.002)
TL/TA	0.960	1.088	0.591	0.629	0.201	0.734	0.168	0.734	0.317	0.146	0.480
	(0.082)	(0.035)	(0.042)	(0.016)	(0.028)	(0.046)	(0.055)	(0.031)	(0.024)	(0.021)	(0.008)
LA/TL	-0.327	-0.476	-0.478	-0.371	-0.091	-0.190	-0.168	-0.299	-0.233	0.011	-0.237
	(0.093)	(0.040)	(0.042)	(0.020)	(0.040)	(0.036)	(0.049)	(0.033)	(0.017)	(0.009)	(0.008)
I/TS	0.021	0.323	-0.207	0.264	1.310	0.206	0.008	0.047	0.297	0.067	0.115
	(0.050)	(0.036)	(0.044)	(0.017)	(0.310)	(0.240)	(0.056)	(0.009)	(0.041)	(0.021)	(0.006)
TL/TS	0.167	0.124	0.301	0.148	0.224	0.082	0.040	0.064	0.198	0.353	0.162
	(0.025)	(0.006)	(0.008)	(0.004)	(0.013)	(0.011)	(0.015)	(0.006)	(0.006)	(0.006)	(0.002)
IP/(IP+EBITDA)	0.089	0.092	0.048	0.054	-0.002	0.174	0.054	0.157	0.039	0.138	0.079
	(0.037)	(0.013)	(0.014)	(0.007)	(0.019)	(0.025)	(0.053)	(0.017)	(0.012)	(0.014)	(0.004)
PAYREMARK	1.449	1.604	1.854	1.643	1.616	1.815	2.369	1.773	1.894	2.592	1.838
	(0.125)	(0.046)	(0.046)	(0.028)	(0.059)	(0.062)	(0.159)	(0.053)	(0.037)	(0.044)	(0.015)
TAXARREARS	2.910	2.361	2.652	2.579	2.468	2.951	3.210	2.538	2.997	2.786	2.698
	(0.081)	(0.029)	(0.030)	(0.018)	(0.041)	(0.042)	(0.112)	(0.034)	(0.025)	(0.027)	(0.010)
PAYDIV	-2.168	-1.674	-1.627	-1.922	-1.493	-1.549	-0.977	-1.444	-1.579	-2.077	-1.809
	(0.400)	(0.114)	(0.111)	(0.076)	(0.219)	(0.179)	(0.296)	(0.140)	(0.085)	(0.133)	(0.039)
TTLFS	4.070	3.593	3.941	3.551	3.278	3.864	3.680	3.460	3.720	3.300	3.587
	(0.073)	(0.027)	(0.029)	(0.016)	(0.040)	(0.039)	(0.097)	(0.030)	(0.022)	(0.022)	(0.009)
Aggregate variables [b]											
Output gap	-0.128	-0.120	-0.156	-0.104	-0.111	-0.126	-0.129	-0.187	-0.120	-0.040	-0.115
	(0.020)	(0.007)	(0.007)	(0.004)	(0.010)	(0.010)	(0.029)	(0.008)	(0.006)	(0.006)	(0.002)
Nominal interest rate	0.058	0.072	0.088	0.073	0.048	0.050	0.093	0.082	0.073	0.060	0.072
	(0.015)	(0.005)	(0.006)	(0.003)	(0.008)	(0.008)	(0.021)	(0.006)	(0.005)	(0.005)	(0.002)
GDP inflation	-0.022	0.014	-0.034	0.016	0.036	0.024	-0.053	-0.013	0.006	0.011	0.006
	(0.021)	(0.007)	(0.008)	(0.005)	(0.012)	(0.012)	(0.033)	(0.009)	(0.007)	(0.008)	(0.003)
Real exchange rate	0.000	-0.011	-0.002	-0.003	0.000	-0.010	-0.011	-0.007	-0.008	-0.009	-0.006
	(0.005)	(0.002)	(0.002)	(0.001)	(0.002)	(0.002)	(0.007)	(0.002)	(0.001)	(0.001)	(0.001)
Mean log-likelihood	-0.024	-0.043	-0.043	-0.050	-0.070	-0.035	-0.025	-0.059	-0.031	-0.058	-0.045
Pseudo R^2.	0.40	0.30	0.38	0.33	0.31	0.39	0.42	0.35	0.39	0.41	0.35
Pseudo R^2. | agg.coeffs. [c]	0.36	0.29	0.37	0.32	0.30	0.39	0.34	0.34	0.38	0.39	0.35
Industry R^2	0.88	0.95	0.95	0.97	0.85	0.89	0.84	0.86	0.94	0.83	0.96
Industry R^2 | agg.coeffs. [c]	-2.01	0.87	0.89	0.90	0.63	0.71	-1.82	0.78	0.34	0.55	0.96
Number of obs	219,604	1,035,359	848,658	2,161,664	251,693	515,674	99,363	473,831	1,739,890	760,402	8,106,138

Notes: Standard errors in parentheses. The variables are not scaled, so the importance of a variable cannot be interpreted directly from the size of the parameter estimate. [a] See Subsection 2.1 for exact definition of these variables. [b] See Subsection 2.2 for definition and sources. [c] Pseudo R^2. | agg.coeffs. is the Pseudo R^2 value calculated for each industry using the estimated coefficients in the economy-wide model (i.e., the coefficients in the last column in the table above). The pseudo R^2 values are calculated according to McFadden (1974).

Table 4: Out-of-Sample Root Mean Squared Error (RMSE) for various models

Model	RMSE (in percent) [b]											
Absolute RMSE for Model j [a]	**Agriculture**	**Manu-facturing**	**Construction**	**Retail**	**Hotel & Restaurant**	**Transport**	**Bank, Finance & Insurance**	**Real-Estate**	**Consulting & Rental**	**Not Classified**	**Industry aggregate**	**Economy Wide**
Only firm-specific variables	0,1973	0,3039	0,4239	0,4509	0,7457	0,2641	0,2070	0,7079	0,2504	0,2680	0,3427	0,3350
Firm-specific and macro	0,0711	0,0849	0,0842	0,1215	0,3210	0,0697	0,1013	0,2459	0,1176	0,2381	0,0660	0,0478
Economy-wide coefficients	0,2830	0,0904	0,1612	0,0540	0,3454	0,0789	0,2124	0,3490	0,1728	0,7155	0,0478	0,0478
Time series random walk	0,1082	0,1179	0,1023	0,1180	0,2338	0,1119	0,1400	0,0737	0,1133	0,3576	0,1262	0,1262
Industry OLS macroregression	0,1216	0,1288	0,2327	0,1788	0,4608	0,1608	0,1081	0,4265	0,9129	0,3934	0,2270	0,1419
4 quarter moving average	0,0854	0,1208	0,0772	0,0797	0,1570	0,0782	0,1137	0,0761	0,0869	1,5321	0,0893	0,0893
RMSE model j / RMSE Table 3 model [a]												
Only firm-specific variables	2,7752	3,5783	5,0368	3,7114	2,3231	3,7896	2,0435	2,8789	2,1291	1,1257	5,1901	7,0059
Economy-wide coefficients	3,9806	1,0646	1,9155	0,4444	1,0759	1,1316	2,0963	1,4191	1,4687	3,0055	0,7243	1,0000
Time series random walk	1,5219	1,3885	1,2152	0,9712	0,7283	1,6053	1,3817	0,2998	0,9636	1,5019	1,9121	2,6399
Industry OLS macroregression	1,7103	1,5171	2,7637	1,4716	1,4355	2,3070	1,0671	1,7344	7,7628	1,6522	3,4392	2,9686
4 quarter moving average	1,2007	1,4230	0,9173	0,6562	0,4891	1,1217	1,1225	0,3096	0,7385	6,4354	1,3522	1,8668

Notes: The RMSEs have been computed as one-step-ahead forecasts for the period 2000Q1-2002Q4. The RMSE ratios have been computed relative to the second row in the first panel, i.e., the industry-specific models. All models were estimated for the period 1990Q1-1999Q4. Industry aggregate RMSEs have been computed by summing the default frequency probabilities implied by each industry model quarterly. [a] Note that the macro variables in these forecasting models are lagged one quarter, so that all models are based on the same information set. [b] Notice that the RMSE numbers are expressed in percent, i.e., fitted and actual default numbers are multiplied by 100 before the RMSE numbers are computed.

Table 5: Out-of-sample Pseudo R^2 and decile tests at the industry level

Industry-specific model coefficients

	Agriculture	Manufacturing	Construction	Retail	Hotel & Restaurant	Transport	Bank. Finance & Insurance	Real-Estate	Consulting & Rental	Not Classified	Industry aggregate
Pseudo R²	0,36	0,31	0,46	0,33	0,34	0,41	0,40	0,32	0,38	0,46	0,40
Decile											
1	0.74	0.71	0.85	0.72	0.76	0.82	0.79	0.75	0.78	0.78	0.79
2	0.11	0.13	0.06	0.10	0.09	0.07	0.03	0.06	0.08	0.08	0.08
3	0.08	0.05	0.03	0.06	0.03	0.05	0.05	0.03	0.05	0.10	0.04
4	0.02	0.05	0.02	0.05	0.02	0.02	0.05	0.04	0.03	0.01	0.03
5	0.02	0.02	0.01	0.03	0.03	0.01	0.04	0.04	0.02	0.01	0.02
6 - 10	0.02	0.04	0.03	0.05	0.07	0.03	0.05	0.08	0.04	0.03	0.04
Sum	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

Economy-wide model coefficients

	Agriculture	Manufacturing	Construction	Retail	Hotel & Restaurant	Transport	Bank. Finance & Insurance	Real-Estate	Consulting & Rental	Not Classified	Aggregate
Pseudo R²	0,35	0,31	0,46	0,33	0,34	0,41	0,35	0,35	0,38	0,46	0,39
Decile											
1	0.75	0.69	0.86	0.71	0.76	0.81	0.76	0.72	0.78	0.78	0.76
2	0.11	0.12	0.05	0.09	0.09	0.08	0.04	0.05	0.08	0.04	0.09
3	0.06	0.07	0.03	0.06	0.02	0.04	0.06	0.06	0.05	0.11	0.06
4	0.03	0.05	0.02	0.05	0.04	0.02	0.04	0.04	0.03	0.03	0.03
5	0.03	0.03	0.01	0.03	0.02	0.01	0.03	0.02	0.02	0.01	0.02
6 – 10	0.02	0.04	0.03	0.05	0.06	0.03	0.08	0.10	0.04	0.02	0.04
Sum	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**	**1.00**

Notes: The out-of-sample period is 2000*Q*1-2002*Q*4, and the total number of firms in the panel for this period is 2,614,248. The decile test outcomes in the table are obtained by sorting the estimated default probabilities, in descending order, and by computing observed default frequencies in the different deciles of the sorted data. The coefficients used for calculating the default probabilities are the ones presented in Table 3. Industry aggregate numbers are obtained by generating the estimated default probabilities using industry-specific model coefficients, then aggregating observations over the various industries to form a single data set, to which, finally, the procedure outlined above is applied in order to compute default frequencies for various deciles.

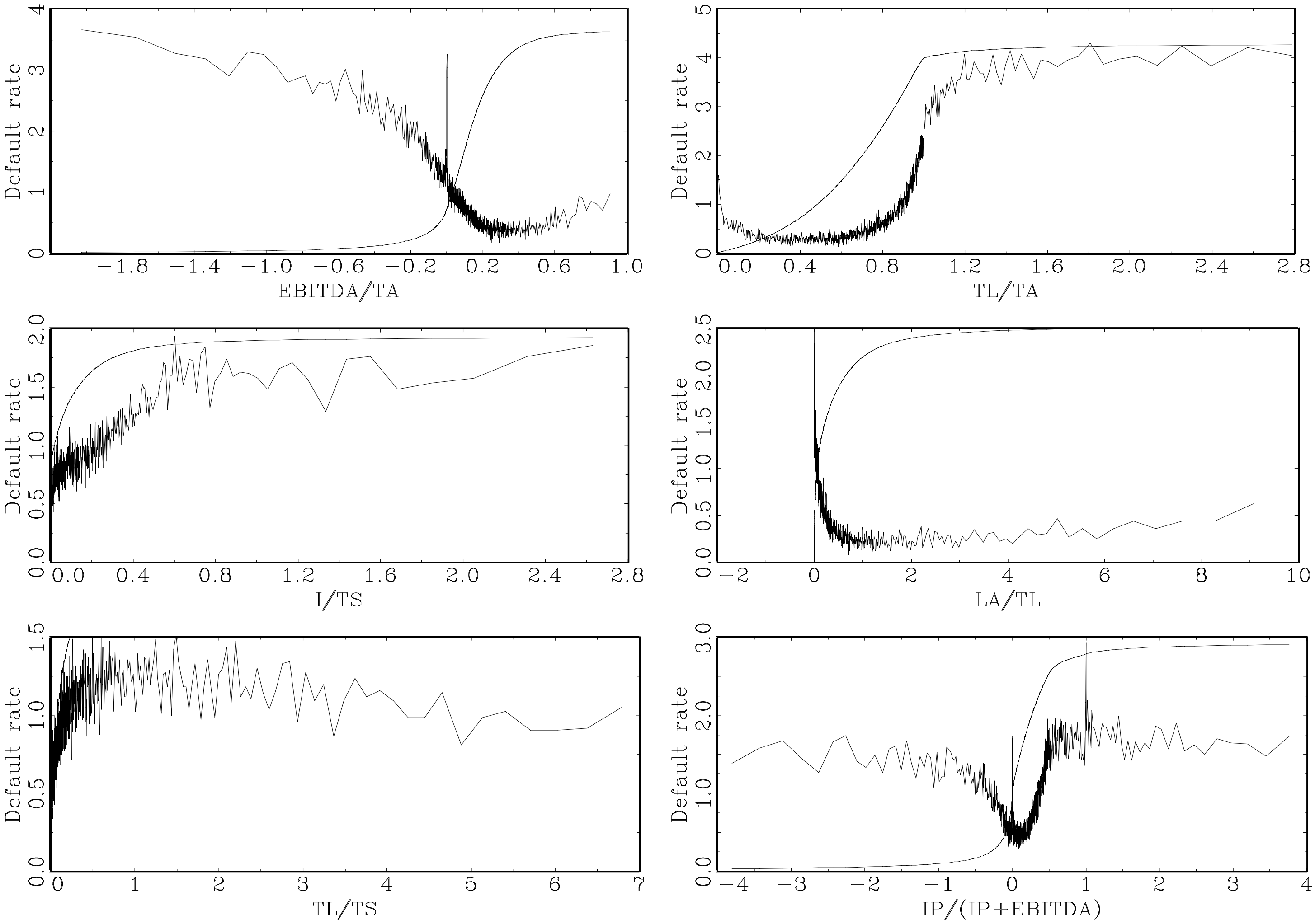

Figure 1: Default rates and the cumulative distribution functions for the accounting data.

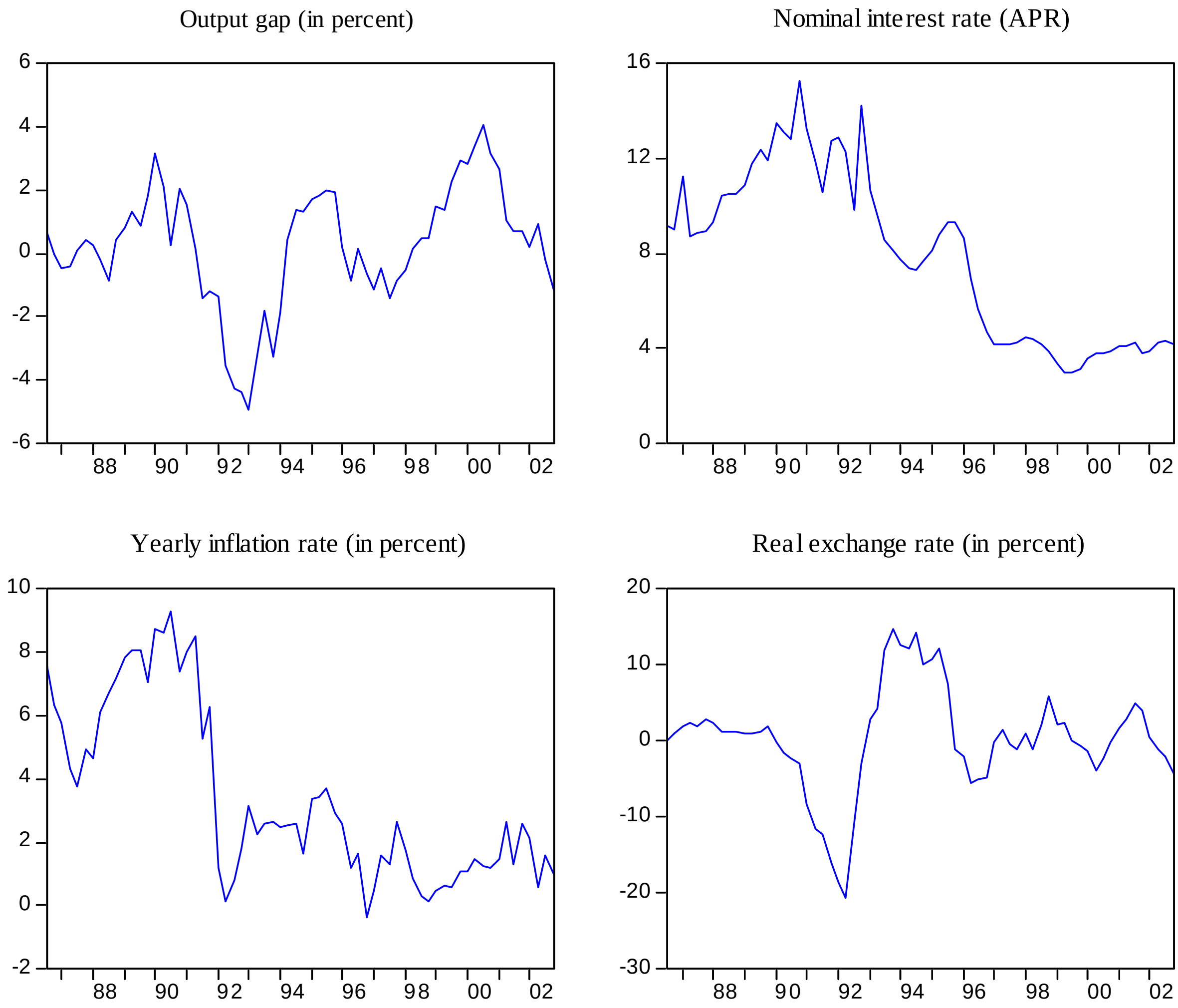

Figure 2: Macro data used in the estimated models.

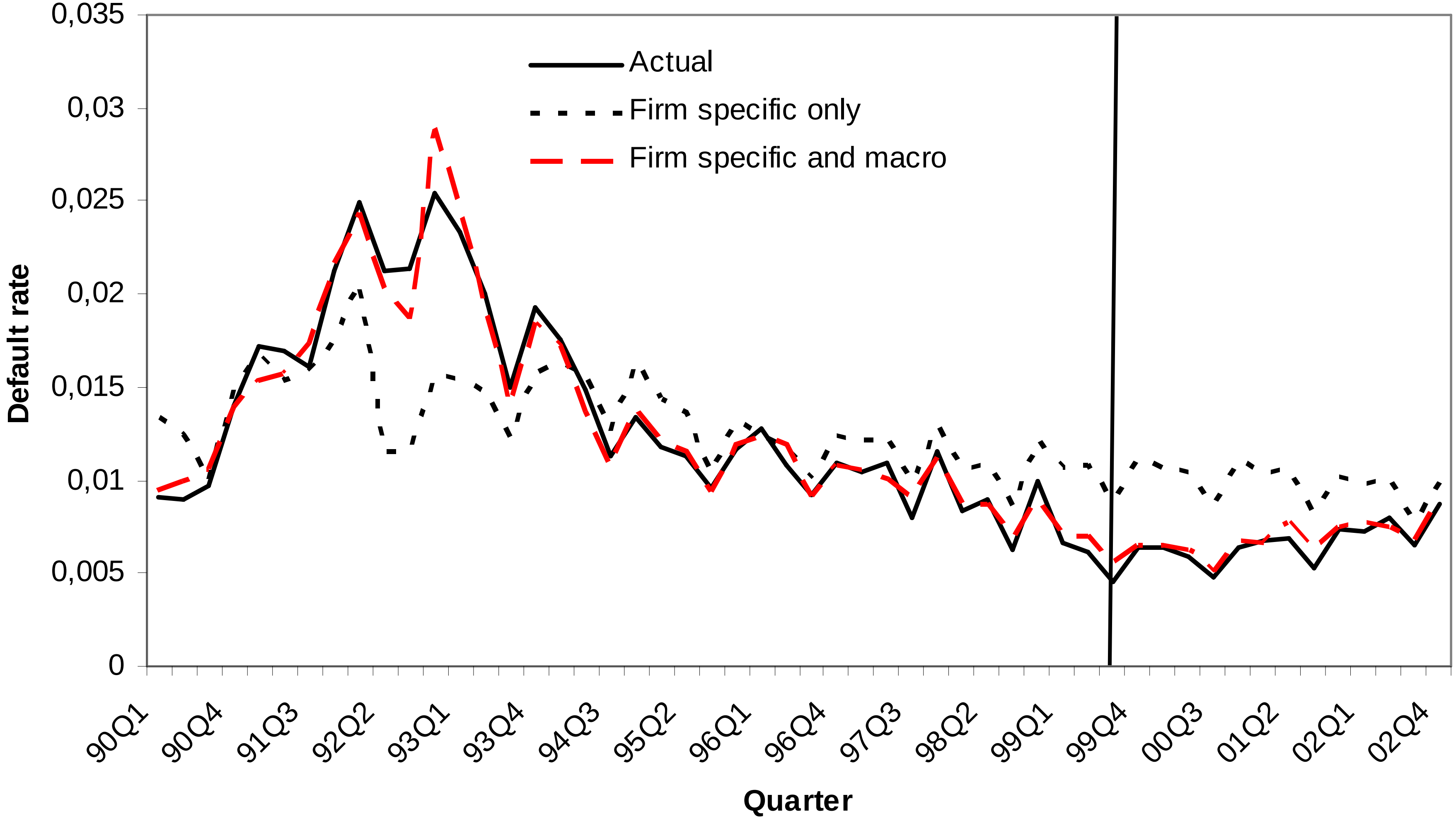

Figure 3: : Actual (solid) and projected (dashed, dotted) aggregate default frequency rates 1990Q1-2002Q4. The projected rates are constructed using the estimated economy-wide models in Table 2 (dotted) and Table 3 (dashed). The models are estimated on data until 1999Q4. The projections shown to the right of the vertical line are out-of-sample.

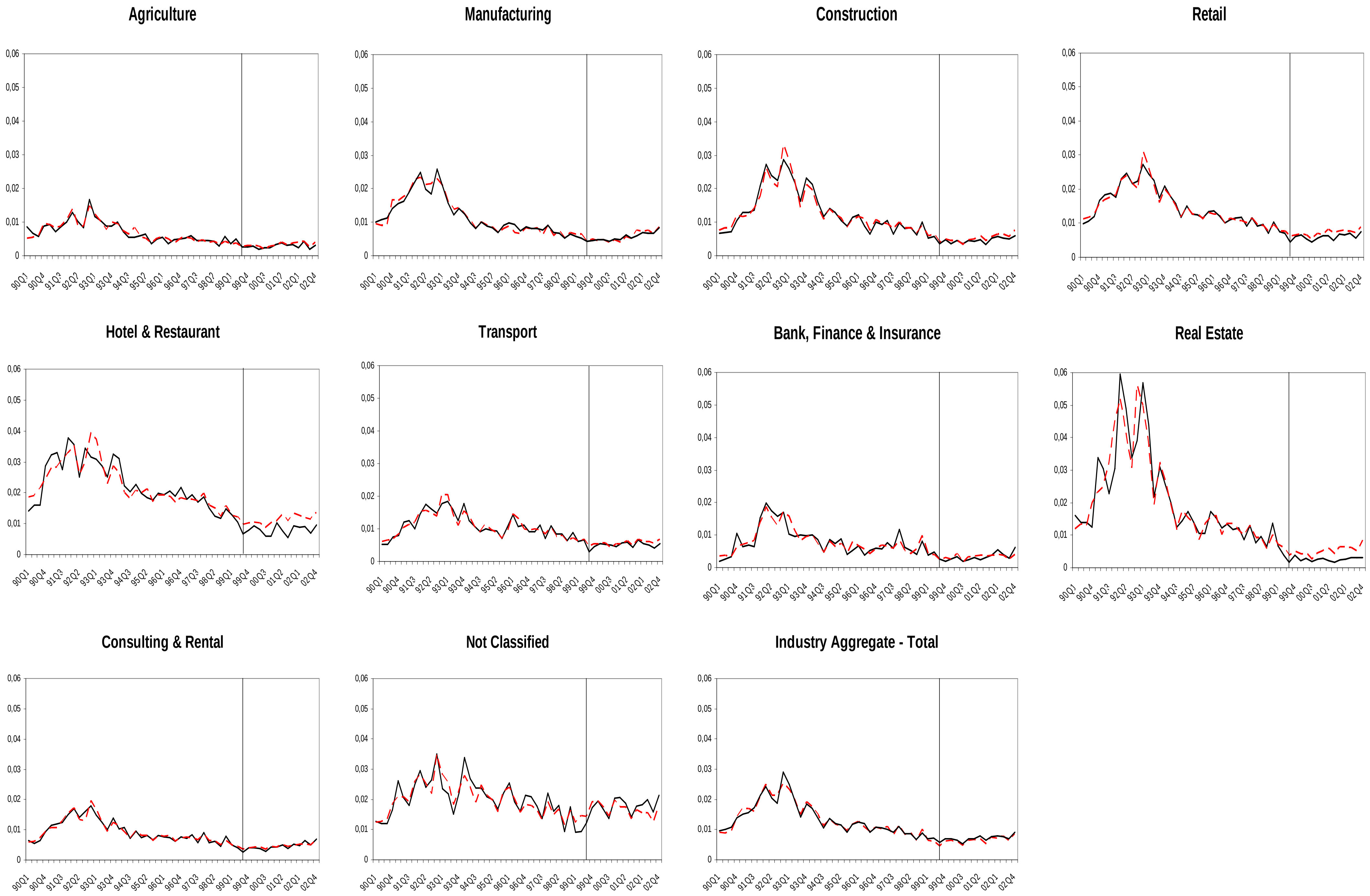

Figure 4: Actual (solid) and projected (dashed) industry default frequency rates 1990Q1-2002Q4. The projected rates are constructed using the estimated industry-specific models in Table 3. The models are estimated on data until 1999Q4. The projections shown to the right of the vertical line are out-of-sample.

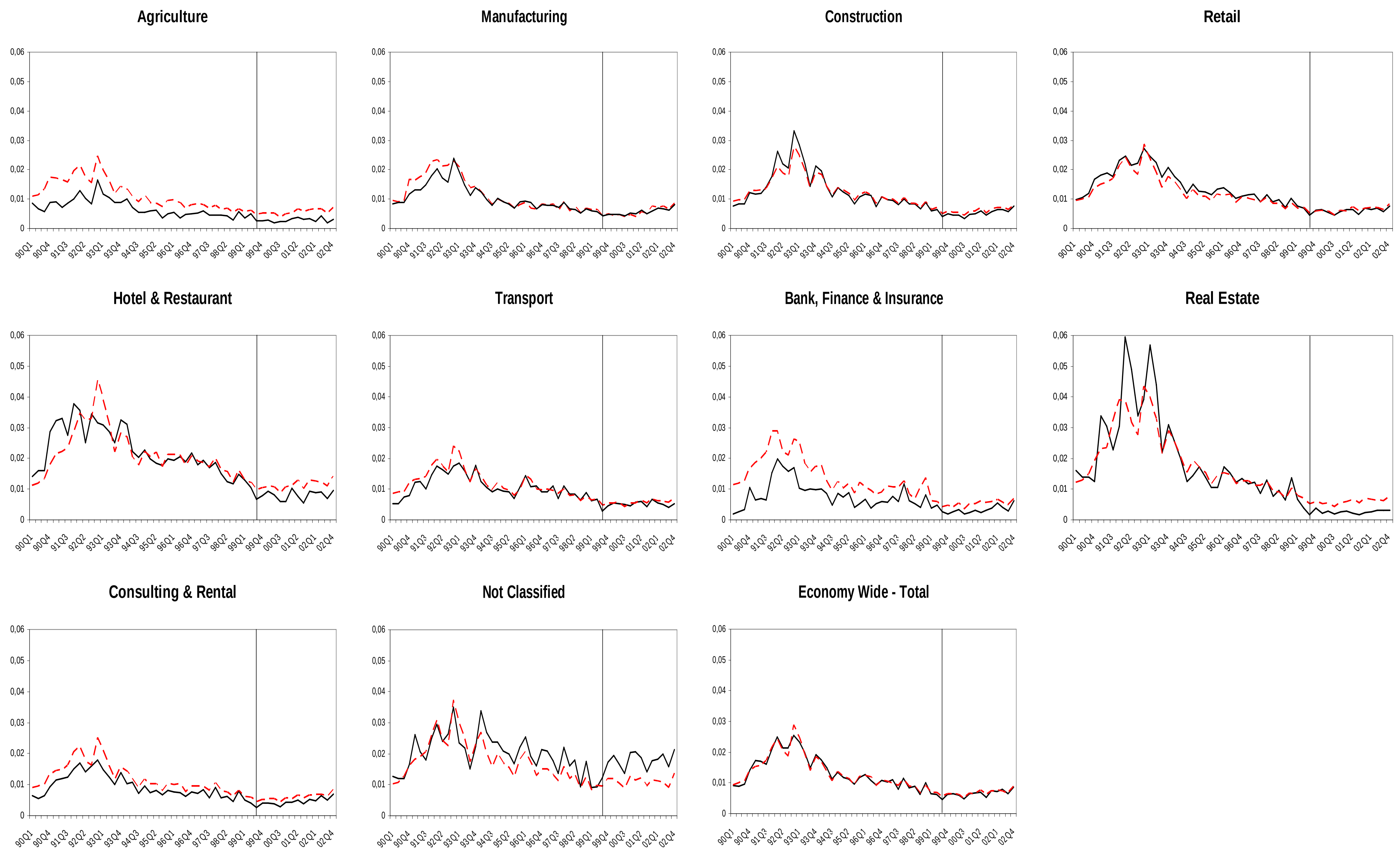

Figure 5: Actual (solid) and projected (dashed) industry default frequency rates 1990Q1-2002Q4. The projected rates are constructed using the estimated economy-wide model in Table 3. The model is estimated on data until 1999Q4; hence the projections to the right of the vertical line are out-of-sample.

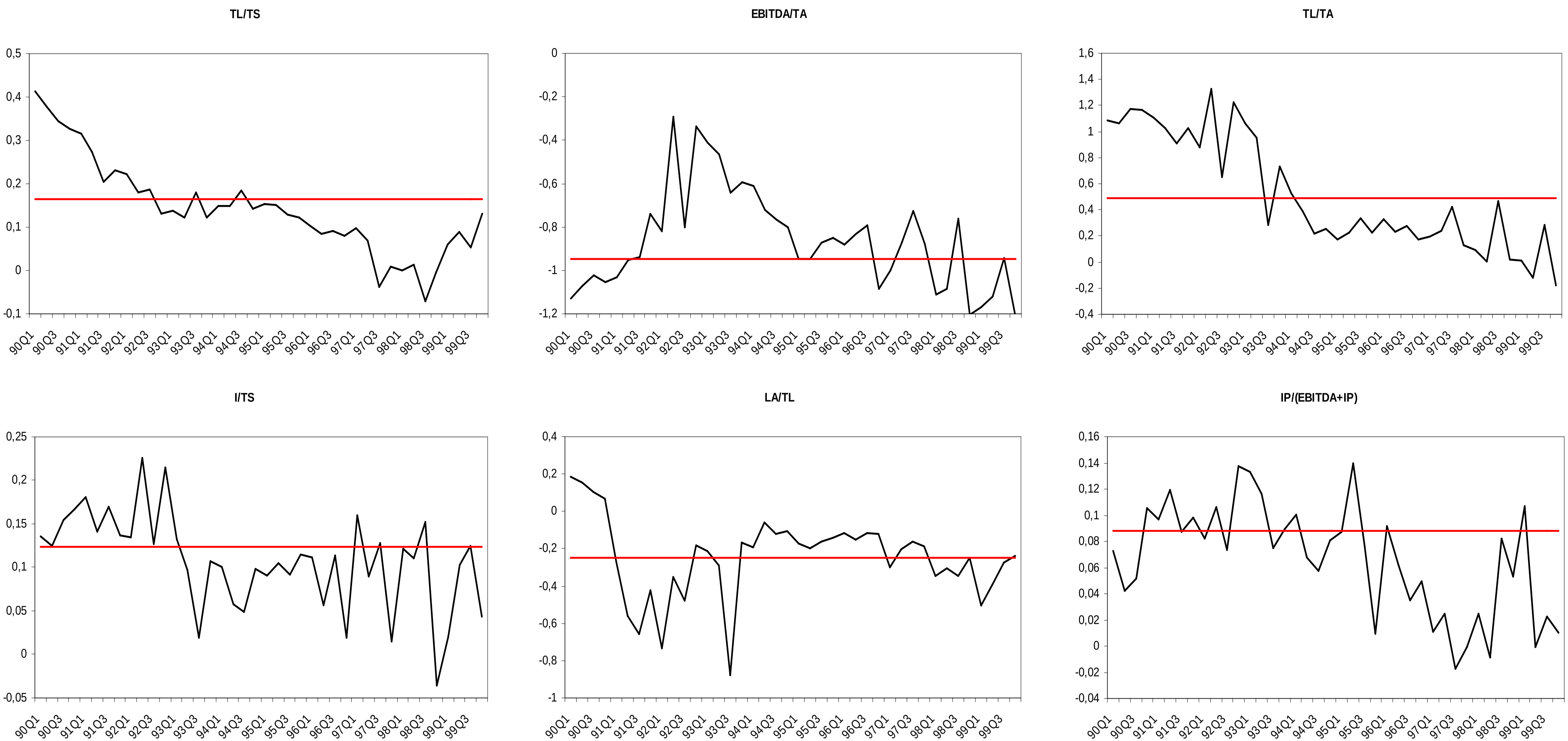

Figure 6: Time-varying coefficients for the accounting ratios in the economy-wide model estimated in each quarter without the macro variables included up to 1999Q4. The horizontal lines correspond to the estimated coefficients in the economy-wide model in Table 2.

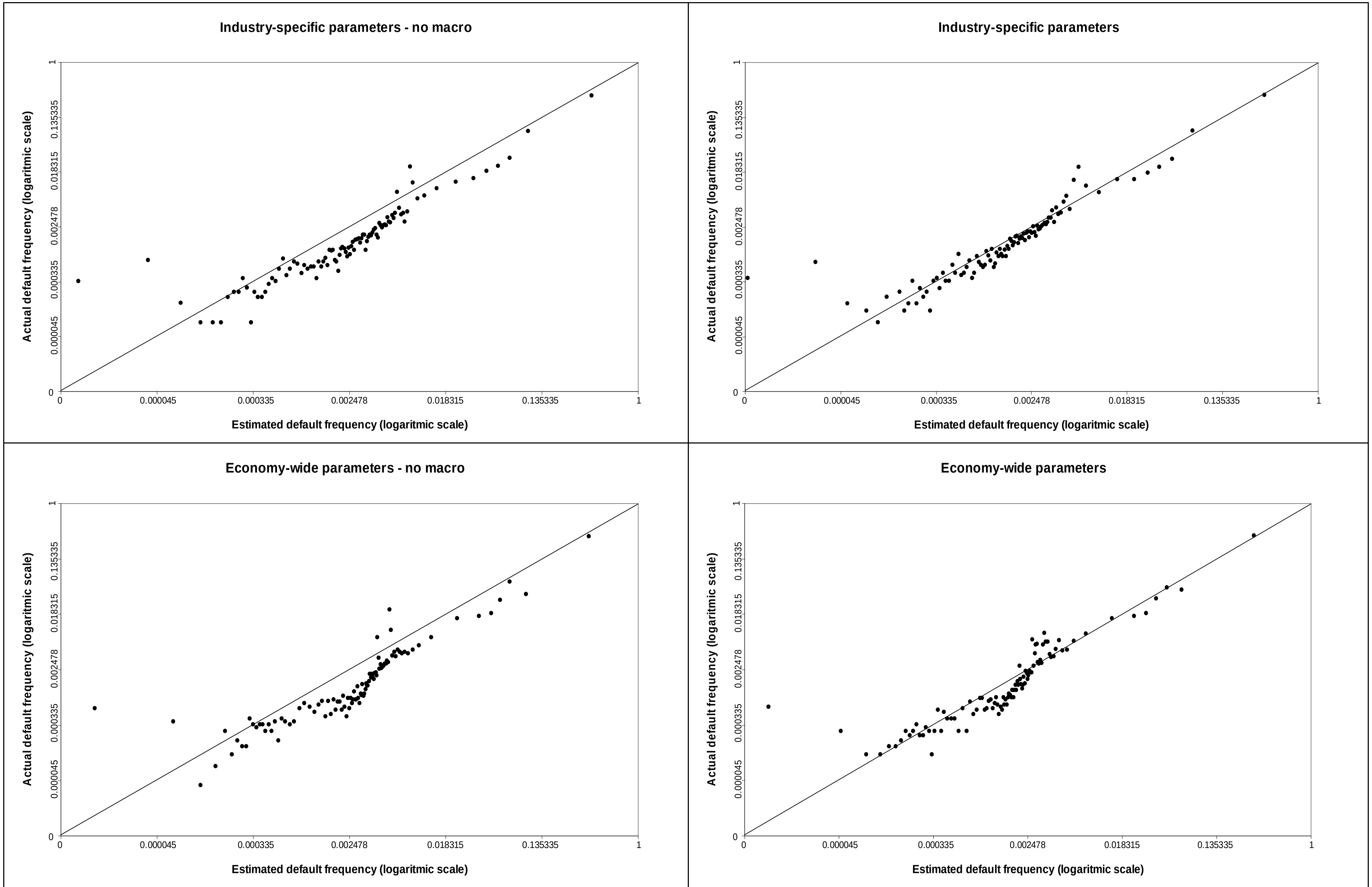

Figure 7: Sorted estimated default percentiles versus actual default frequencies for both economy-wide and industry-specific parameters. Left panel: Only firm-specific variables included (Table 2 models); Right panel: Macro variables included (Table 3 models).