Aggregate shocks and Firm Default Risk *

Tor Jacobson Jesper Lindé Rikard Nilsson Kasper Roszbach

January 14, 2008

Abstract

In this paper we use a large panel data set to examine if there is a separate role for aggregate shocks on firm default probabilities, conditional on an extensive set of firm-specific factors. We find strong evidence for substantial spillover effects of aggregate shocks on firm default probabilities. In addition, the effects of aggregate shocks are more prominent in some industries than others in a natural way. Given these result, we provide some suggestions as to why aggregate shocks could have an affect on firm default probabilities over and above the influence of firm-specific variables. Finally, we evaluate the properties of the estimated default risk model out-of-sample and find that our estimated model is clearly superior to best fitting naive forecasting models. We also document that the firm-specific factors provide very useful relative riskiness ranking, whereas the aggregate shocks is the most important determinant of the absolute riskiness level. Together, both sets of factors enable the estimated models to make accurate absolute risk-rankings, both in-sample and out-of-sample.

Keywords: Default-risk model; Business Cycles; Micro-data; Firm-specific variables; Macroeconomic variables

JEL: C41, G21, G33, G38

*Research Department, Sveriges Riksbank, 103 37 Stockholm. The views expressed in this paper are solely the responsibility of the authors and should not be interpreted as reflecting the views of the Executive Board of Sveriges Riksbank. We are indebted to Erik von Schedvin for providing excellent research assistance.

1 Introduction

In this paper we study empirically the effect of aggregate shocks on firm-default probabilities. To our disposal, we have a unique dataset covering all Swedish firms limited by shares $1990Q1-2002Q4$. In the empirical analysis, we adopt a very simple econometric specification and use a standard logistic regression to model default risk at the firm level where in addition to an extensive set of firm-specific variables, a set of standard macroeconomic variables are considered. The logistic regression model have been applied in earlier studies on firm default, by e.g. Altman and Saunders (1997), Shumway (2001) and Jacobson, Lindé and Roszbach (2005). The default risk models are estimated on the sub-sample $1999Q1 - 1999Q4$, and are then used to assess the estimated models in-of-sample fit along with a thorough examination of the out-of-sample performance $2000Q1 - 2002Q4$. Because the aggregate default frequency and macroeconomic outcome have been very volatile during the 1990s, we think of the out-of-sample exercise as an important step in order to assess to what extent our model can be viewed as causal, and hence supporting our hypothesis that aggregate shocks are important for understanding default behavior at the firm level over and above the extensive set of firm-specific factor that are included in the empirical model. Since we have as many as $250,000$ firms in the panel each quarter, we are also able to estimate separate models for each industry. In our view, this is an important step in the analysis as one reason why aggregate variables may come out as very important in an estimated logit-model at the aggregate level is that they capture systematic industry differences in firm-specific variables. Therefore, the estimation evidence for the industry specific models can more credibly be used to assess the robustness of the role of aggregate shocks in firm default behavior. Are aggregate shocks more important for some that others in a way that is to be expected; i.e. do we find the evidence of stronger links to the real-estate and construction and real estate industries compared to e.g. the agriculture sector.

The main findings are as follows. First, we find that macroeconomic variables are important for explaining the time-varying default frequency in-of-sample. Firm-specific variables are very useful in ranking the riskiness of firms, but macroeconomic variables are of crucial importance for explaining changes in absolute default risk. Moreover, we find that the estimated default risk models perform very well out-of-sample, both at the aggregate and industry level as well as the microeconomic level. In addition, the effects of aggregate shocks in the industry specific models in-of-sample came out in such a way that was to be expected; demand and interest rate conditions have a strong impact on the Construction and Real-Estate sectors while the dependence of the Agricultural sector on the macroeconomic stance turned out to be weak. Out-of-sample, the industry specific models have a clear edge of a single aggregate model in a way that was to be expected as well, suggesting that the favorable in-of-sample estimation results for the industry specific models are not driven by over-parameterization. By and large, we think these findings are of great interest, as they imply that the Swedish banking crisis in the beginning of the 1990s was not the outcome of shocks that we cannot learn anything from.

Our empirical findings begs the question as to why aggregate variables should have play a separate role in default risk models, why is it not sufficient to condition on firm-specific factors? For instance, one could imagine that there would be no further information in the output-gap variable for predicting default, over and above that given by firms' total sales as a reflection of variation in aggregate demand. In our view, the following two mechanisms offer explanations. First, if it is costly for banks to monitor borrowers, banks may use aggregate information to assess the probability of getting repayment on granted loans. That is, banks may form their credit granting policies on the basis of macro economic 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. Second, and along the same line of reasoning, if the entrepreneur has imperfect information about his own future business prospects, he may resort to using aggregate conditions as a basis for his decision if it is worthwhile to invest more effort in a firm, or declare bankruptcy.

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 aggregate model, and make an assessment of the in-of-sample fit of the estimated models. In Section 4, we make 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. Of particular interest in this section is whether the models' properties in-sample also hold out-of-sample and to what extent there are gains in using the industry-specific models instead of the economy-wide model estimated on all observations in all industries jointly (i.e., implicitly assuming that industry-specific effects are irrelevant). Finally, Section 5 concludes.

2 Data

2.1 Micro data

In this subsection, we provide a detailed description of our data set at the firm level.

The final data set is a panel consisting of $10,720,386$ quarterly observations on incorporated firms, covering ten years of quarterly data for all Swedish *aktiebolag* companies that have issued a financial statement between January 1, 1990, and December 31, 2002. *Aktiebolag* are by approximation the Swedish equivalent of US corporations and UK limited businesses. Swedish law requires every *aktiebolag* to have at least SEK $100,000$ (approximately US\$ $10,000$) of equity to be eligible for registration at the Swedish Patent and Registration Office (PRV). Firms are also required to submit an annual report to PRV. Small firms such as general partnerships, limited partnerships and sole proprietors will be disregarded, since, as reported by Jacobson and Lindé (2000), incorporated firms by far account for the largest fraction of loans and, also, display the most cyclical variation in default risk.

The firm-data come from Upplysningscentralen AB (UC), a major credit bureau in Sweden, and are from two general sources of information. First, UC has provided us with balance-sheet and income statement data from the firms' annual reports submitted to PRV. These annual report data cover the period January 1, 1989 to December 31, 2002. Second, UC has provided us with historical data on events related to payment remarks and payment behavior for the firms and for their principals. The UC-data are available at different frequencies, varying from daily for payment remarks to (most often) annually for accounting data. We will discuss the specifics of the data in greater detail below.

The accounting data contains information on most standard balance-sheet and income statement variables. In addition to the annual report data, we have information on the firms' track records regarding payment behavior as recorded by payment remarks for 61 different credit and tax related events. The storage and usage of payment remarks are regulated by the Credit Information Act, the Personal Data Act and overseen by the Swedish Data Inspection Board. Examples of events that are registered are: delays in tax payments, the repossession of delivered goods, the seizure of property, the resettlement of loans and actual bankruptcy. In practice, with a record of payment remarks individuals will not be granted any new loans and businesses can find it very difficult to open new lines of credit.

We define the population of existing firms in quarter t as the firms which have issued a financial statement covering that quarter and are classified as "active". For a firm to be classified as active, we require that is has total sales and total assets over $1,000$ SEK (roughly US$ 100). In addition to these firms, we add the firms which according to the data set on remarks are classified as defaulted firms.[1] The adopted definition of default is the one used by UC.[2]

In Table 1, we report the means and standard deviations for the employed accounting ratios and other variables, such as payment remarks and average delayed time to the last issued financial report for the defaulted and non-defaulted firms at the aggregate and industry level for the in-of-sample period $1990Q1 - 1999Q4$. The reason why we restrict these statistics to pertain for this sub-sample only is that we have developed the model on this sub-sample, and the main model evaluation is then conducted for the out-of-sample period $2000Q1 - 2002Q4$. Moreover, the industry analysis will be conducted at the one-digit level only, in order to assure that sufficiently many default observations are included in each industry along both the cross-section and time series dimension. Because of varying availability of data, the statistics in Table 1 are calculated based on slightly different numbers of observations among the variables in a given industry. As indicated by the large standard errors in Panel A of Table 1, showing non-truncated data, there are some accounting data observations that clearly are severe outliers. These observations would seriously distort the estimation results if they were to be included in the logit model. Therefore, we have truncated the top and bottom 1 percent observations for the accounting variables in each industry.[3] Given the large number of observations, this approach is more or less equivalent to simply deleting 1 percent of the observations that have accounting data that fall outside a certain region. Notice that we choose to truncate the observations in each industry rather than at the aggregate level, thereby implicitly allowing for some more dispersion and different means in different industries. Panel B of Table 1 shows the descriptive statistics for the truncated

[1] There is a simple reason why we need to add firms that have defaulted to the population of firms defined by the accounting data. Many firms that default choose not to submit their compulsary annual reports in the year, or even years, prior to default. Hence, the only records of their existence that we have come from the remark registers.

[2] According to the UC-definition, a firm has default status if any of the following more important events have occured: the firm is declared bankrupt in the legal sense, it has suspended payments, it has negotiated a composition settlement, it is undergoing a re-construction, or, distraint with no assets. We differ somewhat from the credit bureau's definition though, in that we use a one quarter horizon, whereas they currently employ a one-year horizon.

[3] This approach is quite common in the literature, and e.g., Shumway (2001) also truncate 1 percent of the top and bottom observations. It should be emphazied that the results are not at all sensitive to varying the truncation rate between 0.5 and 2 percent.

micro data set.[4]

Before we decided to restrict our attention to the set of financial ratios that are shown in Table 1, we studied a number of commonly used accounting ratios that were employed in frequently cited articles studying bankruptcy risk and the balance-sheet channel, but the ones reported show the strongest correlation with default risk.[5] In the analysis, we employ six accounting ratios: earnings before interest, depreciation, taxes and amortization over total assets (earnings ratio); interest payments over the sum of interest payments and earnings before interest, depreciation, taxes and amortization (interest coverage ratio); total liabilities over total assets and total liabilities over total sales (debt ratios); cash in relation to total liabilities (cash ratio); and inventories over total sales (turnover ratio).[6] 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 deleted 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 risk 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-risk 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; low (negative earnings) can turn this ratio highly negative if interest costs are high but earnings are slightly more negative, and this event is naturally associated with an increased default risk. On the other hand, high interest payments and low earnings will also

[4] From Table 1, comparison of the descriptive statistics for the untruncated data makes it clear that defaulted firms are unproportionally more affected when truncating all the observations simultaneously. Since the REMARK1, REMARK2, PAYDIV and TTLFS are dummy variables that are unaffected by our truncation procedure, it may 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 truncation suggested a somewhat larger role for the accounting ratios, but the over-all picture remains the same.

[5] See Altman (1969, 1971, 1973, 1984), Carling et al. (2004), Frydman, Altman and Kao (1985), and Shumway (2001).

[6] 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, appears to contain predictive power for default risk. We therefore decided to include $\mathrm{TL}_{i,t}$ as separate variable, but scaled it with average total sales in period t to obtain a stationary accounting ratio. So the debt to sales ratio is actually defined as $\mathrm{TL}_{i,t}/\mathrm{TS}_t$,where TS_t denotes average total sales in period t.

make this ratio large, and is likewise associated with an increased default risk. Similar reasoning can be be applied to the other ratios as well. What is important to note is that this feature for some of the financial ratios does not imply that these variables are uninformative for default risk in the empirical model. The reason for this being that the correlations between these financial ratios in the cross section are substantial, which makes each of these variables contribute to predicting default risk in the joint empirical model.[7] 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 whether a firm has paid out dividends or not. We therefore included this information as a dummy variable (PAYDIV) in the model, taking the value of 1 if the firm has paid out dividends and 0 otherwise.

As mentioned previously, some firms classified as active or defaulted have not submitted a financial report for every quarter, so there is a missing explanatory variables problem. Rather than excluding these firms from the sample, we decided to replace missing values by the panel mean for the joint set of defaulted/non-defaulted firms for firms where accounting data are not available. Because of this, we also included a dummy variable, denoted TTLFS, which equals unity if a firm has not issued a financial statement one and a half year prior to default, and zero otherwise.[8] The reason for including this variable in the default-risk model is the notion that firms who are about to default are less willing to report information about their financial status. By comparing defaulting and healthy firms in Table 1 we see that this mechanism is at work in the panel.

There is some, but not huge, variation in the average accounting 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 recieved payment remarks are around 0.15 and 0.45, respectively, wheras corresponding shares for non-defaulted firms are 0.00 and 0.03. The Hotel & Restaurant-industry is the outlier. Hence, these firms have the lowest earnings ratios,

[7] 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.

[8] There are 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 in year $\tau + 1$. By letting this dummy variable equal unity with a 6 quarter time lag we do take the real-world time delay into account. Second, given the way we define the population of existing firms, firms that are newly registered and enter into the panel would automatically be assigned TTLFS $= 1$ in the third quarter of their existence since they have not 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 statements. Third, for defaulting firms that are in the panel but have never reported any accounting data prior to default, 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.

largest debt ratios, greatest occurences of payment remarks and least of dividend payments to shareholders, and as consequence, the largest default rate over all.

For the remark variables, we employ the same approach as in Carling et al. (2004) 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 as much as possible and (ii) are highly correlated with default. As it turned out, many remark variables are either contemporaneously correlated with default or lack a significant correlation with default behavior. For our final model, we constructed the REMARK1-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", and the REMARK2-variable reflects if 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 we consider do not automatically imply a subsequent default incident, so there are no tautological issues involved with the usage of these variables.

2.2 Macro data

The macro data set used in this paper is adopted from Lindé (2002). The variables under consideration are the output-gap (i.e., deviation of GDP around 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, controlled by the Riksbank), and the real exchange rate.[9] Because there is a strong trend for the real exchange rate during the sample period, this variable is detrended as well.[10] The aggregate time series are depicted in Figure 2. Since Sweden is an open economy, one might reasonably consider it to be important to condition on foreign variables in the empirical analysis as well. Our results suggest that while foreign variables are an important source of fluctuations in the considered macro variables (again, see Lindé, 2002), it is not necessary to condition directly on foreign variables in the default risk model given that the above mentioned domestic variables are included.

[9] 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.

[10] Lindé (2002) estimates a VAR with 2 lags for the period $1986Q3 - 2002Q4$ and generates a trend for the variables by doing 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 The estimated default-risk models: In-of-sample fit

In this section, we examine if default risk at the firm level is affected by aggregate shocks over and above firm-specific information. We also study the in-of-sample gains of estimating separate models for each industry, and assess the role of aggregate shocks for improving the models fit. The in-of-sample period is chosen to be $1990Q1 - 1999Q4$. The reason for choosing this period as the in-of-sample is that we originally developed the default risk model for this period on aggregate data, without having access to neither subsequent data nor a consistent industry classification over time, see Jacobson, Lindé and Roszbach (2005).

3.1 The default-risk models

In this subsection we present a reduced form statistical model for estimation of probability of default for all Swedish incorporated firms. The model specification is very similar to the Maximum likelihood Logit approach used in Jacobson, Lindé and Roszbach (2005). Thus we propose to estimate the following model:

$$y_i^{\tau} = x_i(\tau)\beta + \varepsilon_i^{\tau},$$

$$\text{where} \quad y_i^{\tau} = \begin{cases} 1 \text{ if } x_i(\tau)\beta + \varepsilon_i^{\tau} \geq 0 \text{ (firm defaults)} \\ 0 \text{ if } x_i(\tau)\beta + \varepsilon_i^{\tau} < 0 \text{ (firm stays in business)} \end{cases},$$

under the assumption that $x_i(\tau)$ and ε_i^{τ} are stochastically independent and also independence of the errors between both firms and time points: $f\left(\varepsilon_i^{\tau}, \varepsilon_j^{\tau}\right) = f\left(\varepsilon_i^{\tau}\right) f\left(\varepsilon_j^{\tau}\right)$ for $i \neq j$ and $f\left(\varepsilon_i^{\tau}, \varepsilon_i^{\tau+l}\right) = f\left(\varepsilon_i^{\tau}\right) f\left(\varepsilon_i^{\tau+l}\right)$ for $l \neq 0$.

So far relatively few empirical studies contain a rigorous analysis of the effects from macroeconomic conditions on default behavior and credit risks at the firm level, see e.g. Carling et al. (2004) for a discussion. The logit model of the default probability that we present in this subsection includes both idiosyncratic and macroeconomic explanatory variables.[11] The empirical reason for testing the potential role of aggregate variables in the model is clear by inspection of Figures 2 and 3. In Figure 3, we plot the mean values of the idiosyncratic financial variables that are used in the model $1990Q1 - 1999Q2$. It is obvious that there are no dramatic changes in the financial ratios during the deep recession 1992-1993. Therefore, a model with only idiosyncratic variables included is unlikely to fully account for the exceptionally high default frequency outcomes at the aggregate level depicted in Figure 2 during $1992 - 1993$. Therefore, we tempted to

[11] For simplicity, we estimate a logit-model rather than a duration model as is done in Carling et al. (2004). Although Carling et al., in contrast with Shumway (2001), found significant evidence of a duration dependence, we believe that this approximation may not be of decisive importance. But an interesting extension of this work is to test for duration dependence in the model.

conjecture that it is important to use aggregate variables in the model. The theoretic arguments why aggregate shocks should be an important determinant of default risk at the aggregate level is perhaps less evident. We have two main arguments why aggregate variables might contain predictive information for firm-default risk over and above the firm-specific information. One candidate hypothesis is that if it very costly for banks to monitor firms, then they are likely to use the macroeconomic stance as an indicator of borrower future repayment ability. Another possibility is that entrepreneur themselves, which have to figure out whether it is worthwhile or not to invest more effort into a given firm, have costs of assessing the future profitability of the firm and may therefore also use the macroeconomic stance as an information device for her decision to declare bankruptcy of not. In addition, if firms are borrowing-constrained, the nominal interest rate will be an important determinant of default risk.

We use standard macro variables in the model, i.e., the output gap, the domestic annual inflation rate, the REPO rate, and the real exchange rate. A priori, we think that these should have a measurable impact on the default risk of any given firm. Starting with the output gap, it may supposedly work as an indicator of demand conditions, i.e. increased demand in the economy reducing default risk. Figure 2 seems, at large, consistent with this view, although there are some spikes in the default rate that presumable have to be attributed to other variables. Also, it is clear from Figure 2 that there has been some variation the output gap around 1996-1998 which has not been met with an increased default rate. Therefore, there are most likely some other aggregate variables that ought to be important as well. Here, we decided to include the nominal interest rate (i.e. the REPO rate) because we know that the nominal interest rate was very high during the recession in the beginning of the 1990s, but has came down substantially after the introduction of the inflation target in Sweden. Given the fact that the export to GDP ratio being around 0.40, the real exchange rate is also a potentially important variable, a deprecation leading to improved competitiveness of Swedish firms. The inflation rate may also be important for firms pricing decisions; higher inflation rates are potentially associated with less certainty about correct relative prices, and may thus lead to potentially higher default risk. Finally, as can be seen from Figure 2, there is a large spike in the REPO rate in the third quarter 1992 due to the fact that the Riksbank raised the so called marginal interest rate to 500 percent unexpectedly and temporarily in order to defend the fixed 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.[12]

In order to document how aggregate variables contribute to default risk, estimation results for two specifications are reported. Table 2 contains results for a model with firm-specific determinants of default risk only (i.e. the accounting ratios augmented with the dummy variables PAYDIV, REMARK and TTLFS), while Table 3 shows results with the macroeconomic variables added.[13] As financial reports issued by firms typically become available 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 to forecast default risk in year $\tau + 1$. To account for this, we have lagged all accounting data by 4 quarters in the estimations. For most firms, who report balance-sheet and income data over calendar years, this means that data for year τ is 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 could is not of any greater importance for the estimated coefficients. When re-estimating the model using contemporaneous data instead, the estimation results were to found to be very similar as the ones reported in Tables 2 and 3.

The results in Tables 2 and 3 show that both idiosyncratic and aggregate information is important for explaining default behavior in both the industry and aggregate models. The variables for omitted (non-reported) financial statements and remarks on firms payment record are the strongest determinants of default in the model. A nice feature of the estimations is that the coefficients for each variable does not change substantially when the model is augmented with the aggregate variables. In particular, the accounting ratios in Table 2 are in general very similar to ones in Table 3. The predictive power of the accounting data appear to be about or slightly more important than the remarks dummies, in particular the indicator variable TTLFS

[12] The estimated dummy coefficient in the VAR equals 28.2 in the REPO rate equation. On the basis of this, we adjusted the REPO rate for this quarter to equal 9.8 percent instead of 38 percent.

[13] In addition to the coefficients reported in Tables 2 and 3, three more variables were included (but not reported). First, a common intercept (a constant). Second, since the bankruptcy rate is systematically lower in the third quarter (most likely due to Swedish courts long vacations in July-August), we included a $0-1$ dummy variable to capture this phenonema. Third, because no data on the payment records of firms (i.e., the dummy variables REMARK1 and REMARK2) exist prior to 1992Q3 for legal storage reasons, the models also include one additional variable (not reported) which is constructed to be an estimate of the average value of the sum of the payment record variables REMARK1 and REMARK2 for the quarters 1990Q1-1992Q2. This variable was constructed by estimating a logit model for the event of either of the dummy variables REMARK1 and REMARK2 taking on the value 0 or 1 for the period 1992Q3-1999Q2, using all the variables in the model in Table 3 (except REMARK1 and REMARK2, of course) as explanatory variables. The imputed average value for this variable for the period 1990Q1-1992Q2 (after 1992Q2, it is set to nil) was then 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 a REMARK1 or a REMARK2 greater than zero, and N_t denotes the number of firms in period t. The largest gain in including this variable is that the effects of the macroeconomic variables in Table 3 are somewhat more accurately estimated. For the coefficients of the idiosyncratic variables this variable is of little importance.

(1 if the firm has not filed a financial report on time and 0 otherwise) and the liability-to-assets ratios (TL/TA and TL/TS) and earnings ratios are quite useful.[14] The turnover ratio for inventories, liquid asset over total liabilities and the interest coverage ratio appear to be less important. We also see that the role of the financial ratios in the various industry model differ substantially; while the role of accounting data is generally low in the Financial services (Bank, Finance and Insurance) sector, it is quite important for default risk in the Manufacturing industry. In the Hotel and Restaurant sector, we find that the I/TS variable is estimated to be very high, whereas it is nil or even negative in the Agriculture and construction industries, respectively. The coefficients for the payment remarks and the indicator variable TTLFS are relatively more similar among the industries, so to the extent that those variables are of rather large importance for explaining firm default behavior, is not clear what is gained at the firm-specific level by conditioning on industry.

Turning to the macroeconomic variables, we find that they are significant in the aggregate model, with the exception of inflation, and have the correct signs. Note that a higher value of the real exchange rate implies a depreciation, and therefore the negative estimate for this variable suggests that a depreciation on average reduces the risk of default at a given point in time. It should be pointed out that the macroeconomic variables are highly significant and quantitatively important even if we allow for non-linear effects of the balance-sheet variables.[15] To our great satisfaction, we see from Table 3 that the impact of the macroeconomic factors are line with was to be expected: Both the output gap and the nominal interest rate are most important in construction and real estate sectors, and the nominal interest rate is quite naturally found to be very important for the Financial services sector as well. The macro variables inflation and the real exchange rate are less important both from a quantitative and a statistical significance perspective, but is comforting that an appreciating real exchange rate (i.e. lower value, see

[14] Regarding the importance of the accounting data in the model, we would like to emphasize the following. As firms typically issue annual financial statements, which we transform into quarterly observations by assuming that they remain the same throughout the reporting period. Given that we define a default event at the quarterly frequency, this assumption could presumably lead to underestimation of the importance of the balance sheet variables in the default risk model. We examined this by estimating the credit risk model at the annual frequency instead, and the coefficients for the balance sheets variables were found to be quite similar. In fact, in the aggregate model, only the coefficients for EBITDA/TA and TL/TS were found to be slightly higher ($-1.2945/0.2652$ instead of $-1.0635/0.1768$, respectively), whereas the other coefficients were actually found to be smaller in the annual model. Similar results were obtained in the industry models.

[15] When estimating the model where the balance-sheet variables enter in a non-linear way (interaction dummies), we used the cumulated distributions depicted in Figure 1 to categorize the balance-sheet 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 resulted in a pseudo R^2 in a non-linear version of the aggregate model in Table 2 of around 0.48, but the aggregate R^2 is still slightly below 0.45. So although this model somewhat better account for the aggregate default frequency, macroeconomic variables are still found to be essential for explaining the absolute level of default risk.

Figure 2) is associated with a significantly lower default risk in the Manufacturing sector, which is arguably the most export-oriented industry.

Finally, we would like to emphasize that the advantage of using firm-specific data when estimating the default-risk model cannot be overstated. If we estimate the model in Table 3 without the dummy variables (REMARK1, REMARK2, PAYDIV, and TTLFS are left out because they do not enter significantly) on aggregate/average data using OLS (TSLS give very similar results), we obtain

$$\begin{aligned} df_t \quad = \quad & \underset{(0.06)}{-0.23} \ \underset{(0.13)}{-0.23} \left(\frac{\text{EBITDA}}{\text{TA}}\right)_t + \underset{(0.06)}{0.30} \left(\frac{\text{TL}}{\text{TA}}\right)_t + \underset{(0.03)}{0.09} \left(\frac{\text{LA}}{\text{TL}}\right)_t \ \cdots \\ & \underset{(0.21)}{-0.94} \left(\frac{\text{I}}{\text{TS}}\right)_t + \underset{(0.08)}{0.19} \left(\frac{\text{TL}}{\text{TS}}\right)_t - \underset{(0.12)}{0.02} \left(\frac{\text{IP}}{\text{IP+EBITDA}}\right)_t \ \cdots \\ & \underset{(0.03)}{-0.05} y_{d,t} - \underset{(0.03)}{0.05} \pi_{d,t} + \underset{(0.03)}{0.12} R_{d,t} + \underset{(0.009)}{0.002} q_t + \hat{u}_{df,t}, \end{aligned} \tag{1}$$

$$R^2 = 0.93, \ \text{DW} = 2.10, \ \text{Sample: } 1990Q1 - 1999Q2 \ (T = 38)$$

If we compare the point estimates in Table 3 with those in (1), we see that they differ substantially. In particular, the balance-sheet variables I/TS and LA/TL account for a lot of the variation in the aggregate default rate, but with the wrong sign. Because the accounting ratios are relatively smooth in the aggregate, which is clear from Figure 3, it is not surprising that we obtain spurious results when estimating the model on aggregate data rather than at the firm level.

3.2 Assessing the models in-of-sample fit

The last rows in Tables 2 and 3 reports the number of observations, mean log-likelihood and the pseudo r-square. The industry r-squares are generated by aggregating all the fitted values in each industry model for each quarter and use the resulting 40 time-series observations to compute the implied r-square. To assess to what extent is it important to estimate industry specific models, we also report the pseudo and industry r-squares using the aggregate parameters instead of the industry estimates.

By comparing Tables 2 and 3, we see that while the pseudo r-squares are not much affected by the conditioning on macroeconomic factors. However, the industry r-squares are about twice as high and sometimes even more than doubled by the introduction of aggregate variables. Thus, the firm-specific variables account for the cross-section of the default distribution, while the macroeconomic stance shifts the mean of the distribution in each period. This also implies that the model with firm-specific information only cannot capture the up- and downturns in average default rate over time visualized in Figure 2, whereas the model with both micro and macro variables included appear indeed able to replicate the high default rate during the banking crisis, 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 rate is 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 high default rates recorded during the banking crisis were not exceptional events that we cannot learn anything useful from, but rather that they were consequences of unusually bad economic outcomes, both domestically and internationally.[16] An additional interesting feature of Tables 2 and 3 is that the fall in pseudo r-square values associated with conditioning on the aggregate parameters is not much pronounced, while the reduction in average r-squares is substantial when imposing the estimates of the aggregate model. This latter results is confirmed in Figure 5, which shows the average industry default frequency along with the projected default frequencies using the aggregate parameter estimates in Table 3. In two cases, for Agriculture, and Bank, Finance & Insurance sectors, we see that the industry r-squares becomes strongly negative conditional on the aggregate parameters. This might seem weird, given that the industry specific parameters in Table 3 are not wildly different. However, as should be clear from Figure 5, these seemly inconsistent results are driven by the not-reported intercept, which is too high in the aggregate model, and thus induces a systematis overprediction of the default risk relative to the actual default risk in these sectors. Given the nature of these sectors, this outcome is not very surprising.

Finally, we would like to take the opportunity and stress the importance of having firm-specific variables in the models. One way of demonstrating how much is lost by omitting the micro structure is to regress the average default frequency on the macroeconomic variables included in Table 3. We then obtain,

$$df_t = \underset{(0.11)}{1.00} - \underset{(0.03)}{0.19} y_{d,t} + \underset{(0.02)}{0.04} \pi_{d,t} + \underset{(0.01)}{0.02} R_{d,t} - \underset{(0.007)}{0.002} q_t + \hat{u}_{df,t},$$

$$R^2 = 0.65,\ \text{DW} = 1.27,\ \text{Sample: } 1990Q1 - 1999Q2\ (T = 38) \tag{2}$$

If we compare the results of this regression with the results in Table 3, we see that we loose about 30 percentage points of the explanatory power in the aggregate model by excluding the balance-

[16] Lindé (2002) shows that a significant portion of the variation in the domestic macroeconomic variables are of foreign origin.

sheet variables. This number is in line with the average R^2 reported for the economy-wide model in Table 2.

4 Out-of-sample properties of the estimated model

In this section, we carefully examine the out-of-sample properties of the estimated models in Table 2, and, in particular, Table 3, for the period $2000Q1-2002Q4$. The number of observations in the panel out-of-sample equals 2,614,248. As already mentioned in Section 3, the reason for choosing this period for our out-of-sample is that the models were originally developed for the in-of-sample period. In the following, we will start out by studying the models properties at the industry and aggregate level, and then turn to the models properties at the firm level.

4.1 Evaluating the models at the industry and aggregate level

In Figures 6 and 7, we extend the content in Figures 4 and 5 with the out-of-sample evidence. We see that in many of the estimated industry models, the fit is remarkably good. 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 have changed. Perhaps not surprisingly, the two sectors are the ones with the highest and most volatile default frequencies during the banking crisis; the Hotel and Restaurant and Real-estate industries. So for these sectors, the transmission of shocks into default behavior appear to have changed. Before drawing to firm conclusions, however, it should also be noted that these sectors contain relatively few observations, so the relatively poor out-of-sample properties of the model for these industries could to some extent be a small sample problem. Admittedly, there are other industries which contain even fewer observations, like Agriculture and Financial services, but those sectors have much more stable default frequency patterns throughout the whole sample period. In general, we also note that the out-of-sample fit is improved by adapting industry specific models compared to the single aggregate model. This will be examined in greater detail below.

In Table 4, we report the root mean squared prediction errors for the models estimated in Tables 2 and 3 along with some standard alternative time series models. The two considered time series model is the classical random walk model, and a simple 4 quarter moving average model. We also experimented with estimated AR(p) models with a dummy for the third quarter included, but they were found to be inferior to the considered models here, in all likelihood due to the rather large differences in the default frequencies between the in- and out-of sample periods.

From inspection of Table 4, we see that the model with aggregate variables included are clearly superior to the model with only firm specific information. The RMSEs for the models in Table 2 have are about $2-3$ times higher. Moreover, the industry specific models generate lower RMSEs compared to the RMSEs generated with the estimates from the aggregate model in Table 3. This is evidence that the industry specific models are not over-parameterized w.r.t. the macro variables. An summary statistic to assess the relevance of the industry specific models at the aggregate level is to compare the RMSE of the industry aggregate to the RMSE for the aggregate model in Table 3.[17] This statistic shows that while both RMSEs are very low, exploiting the industry specific parameters results in an RMSE that is about 20 percent lower than the RMSE using the aggregate estimates. Clearly, this result support the industry specific specifications.

Comparing the Table 3 industry models with the time series models, we also see that the while the random walk is better for 3 out of 10 sectors and the 4 quarter moving average specification is better 6 out of 10 times, they are still both clearly inferior to the industry models at the industry aggregate level. In fact, they are also inferior in terms of RMSE fit to the aggregate model specification in Table 3 (which condition on aggregate fluctuations).

To sum up, we have found strong evidence that the good in-of-sample fit of the estimated industry (and aggregate) models with aggregate shocks also holds well out-of-sample at the sector and aggregate level, suggesting that the macroeconomic factors that enters into the model are structural, and not just improving the in-of-sample fit of the models. We have also documented that there appear to be some gains in terms of forecasting accuracy to condition on the industry specific models rather than just an aggregate model.

4.2 Evaluating the models at the industry and firm level

In this subsection, we examine the out-of-sample properties at microeconomic level. To do this, we follow Shumway (2001) and examine how well the model ranks the firms in terms of default frequencies, i.e. is the riskiest firms assigned the highest default probabilities in the estimated default risk models and vice versa. In addition, we report the pseudo r-squares for the industry

[17] The time series for expected default frequency in the industry aggregate is computed by aggregating all the out-of-sample expected default probabilities generated by the industry specific models in each period.

specific models conditional on the industry specific models in Table 3 as well as the pseudo r-squares calculated with the aggregate parameters. The results are reported in Table 5.

Starting with the pseudo r-squares for the industry specific parameters, we see that they have in fact increased or equal in 7 out of 10 sectors, and in the three cases where they have decreased (Hotel & Restaurant, Transport and Real-Estate) they have done so marginally.[18] Turning to the pseudo r-squares for the aggregate parameters in the lower panel, we see by comparison of Tables 3 and 5 that the pseudo r-squares have been slightly deceased in 3 cases (Transport, Bank, Finance & Insurance, and Real-Estate) and are very similar in the other seven other cases. For the aggregate model, the pseudo r-square have increased slightly (from 0.40 to 0.42). Again, these results strongly suggest that the industry models are not over-parameterized, and that the reduced form parameters have been very stable out-of-sample. Moving to the estimated models ability to rank the relative riskiness of the firms, we first notice that the estimated models roughly classify about $75-80$ per cent of the defaulting firms in the first decile. These numbers are about the same as the ones reported in Shumway (2001) for a substantially smaller data set which only included firms listed on the shock exchange where he was able to condition on market information. Given that our models cover the whole population of firms limited by shares, and only a very small subset of those firms are listed on the stock exchange - about 500 out of $250,000$ - we think our models does a very good job in ranking the firms riskiness. However, Table 5 also reveals that the satisfactory riskiness ranking does not depend on whether we condition on the industry specific parameters or the parameters in the aggregate model. This is in contrast to the findings in the previous subsection, where we found that conditioning on the industry specific parameters substantially improved the models empirical performance. The reason the same conclusion does not apply here is that what is mostly different between the aggregate and industry specific models are the impact of the aggregate factors, and those factors have little impact on the firms' riskiness ranking.[19]

An alternative way to assess the out-of-sample properties of the models in a more absolute sense, is to compare the actual default probability in different percentiles with the average estimated default probabilities in each percentile. To obtain the percentiles, we have added all the out-of-sample observations and sorted them by their estimated default probability and then

[18] Notice that we are not able to compute the pseudo r-square for the industry aggregate using the industry specific parameters, as the Laitila (1993) formulas for computing the r-squares make use of the estimates and their associated covariance matrix.

[19] 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.

computed the average estimated and actual default frequency in each percentile.[20] In Figure 8, we plot the result where we have used both the industry specific and aggregate model parameters in Table 3 to compute the estimate 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, we have taken logs of both the estimated and actual series. If the estimated models could perfectly predict the absolute riskiness of the firms, all dots would lie on the straight line drawn in the figures, which has a slope of unity and intercept equal to nil. As can be seen in Figure 8, this is not the case for neither model, but the dots are generally very close, suggesting that the absolute riskiness ranking are acceptable. Indeed, in a regression

$$Y_i = \beta_0 + \beta_1 \hat{Y}_i + \varepsilon_i,$$

where Y_i is the actual default frequency in percentile $i = 1, ..., 100$ and Y_i is the average estimated default probability in percentile i, the null hypothesis that $\beta_0 = 0$ and $\beta_1 = 1$ cannot be rejected, the $F-$statistics equal 0.04 and 0.02 for the industry and aggregate models, respectively, whereas the critical value at the 10 percent confidence level equals 2.36.

To conclude, both the industry and aggregate models are very accurate in terms of both relative and absolute riskiness ranking at the firm level out-of-sample, in addition to the good empirical performance at the industry and aggregate level documented in the previous subsection.

5 Conclusions [To be completed]

In this paper, we have studied the interaction between the macroeconomic stance and firm default risk at the microeconomic (i.e. firm) level using reduced form methods. To this end we have acquired a large panel data set for the Swedish economy during $1990 - 2002$. A period covering a banking crises episode and associated deep recession in the early 1990s followed by a boom in the latter part of the 1990s, as well as a slight downturn in the beginning of the 2000s. We divided the sample in two parts, $1990 - 1999$ and $2000 - 2002$ and used the former to estimate the models and latter to provide an assessment whether the impact of aggregate fluctuations over and above firm-specific variables are a robust regularity during the whole sample period.

[20] 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 in each industry.

We documented that a simple logit model for default at the firm level using both firm-specific and macroeconomic variables as explanatory variables can explain the extremely high default frequencies during the banking crisis in the beginning of the 1990s, and also the considerably lower default frequencies in the late 1990s. The estimated models were then shown to be very robust and successful out-of-sample, suggesting that aggregate shocks play a truly prominent role in understanding the absolute level of firm default risk. It should be emphasized that we do not want our results to be interpreted to imply that aggregate shocks are the most important source of default at the firm level, rather, our results suggest that macroeconomic factor to be the most important source of average default risk.

References

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

Altman, Edward I., (1971), Railroad bankruptcy propensity, *Journal of Finance*, Vol. XXVI 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.

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.

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. LXXVIII No.1, pp. *1-15.*

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 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, (2004), "Corporate Credit Risk

Modelling and the Macroeconomy", manuscript, Sveriges Riksbank.

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.

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

Frydman, Halina, Edward I. Altman, 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.

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 10*9, 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.

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.

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", manuscript, Sveriges Riksbank.

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.

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.

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, pp. Xx-xx+5.

Table 1: Descriptive statistics for firm-specific micro data 1990*Q*1-1999*Q*4
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	3.7	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)
Dividend	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)
Remarks 1	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)
Remarks 2	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)
Dividend	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)
Remarks 1	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)
Remarks 2	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; REMARK1 = 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; REMARK2 = a dummy variable taking the value of 1 if the firm is in various tax arrears; TTLFS = a dummy variable equal to 1 if the firm has not submitted an annual report in the previous year, and 0 otherwise.

Table 1: Descriptive statistics for firm-specific micro data 1990*Q*1-1999*Q*4
Panel B: 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	-0.04	-0.01	0.01	-0.06	-0.11	0.00	-0.07	0.02	-0.02	-0.08	-0.03
	(0.34)	(0.27)	(0.27)	(0.33)	(0.52)	(0.34)	(0.55)	(0.23)	(0.40)	(0.52)	(0.36)
TL/TA	1.07	0.96	0.95	1.02	1.18	1.01	1.05	1.02	0.93	0.97	1.00
	(0.46)	(0.36)	(0.36)	(0.49)	(0.75)	(0.46)	(0.98)	(0.44)	(0.54)	(0.77)	(0.52)
LA/TL	0.14	0.13	0.16	0.16	0.20	0.20	0.42	0.13	0.32	0.55	0.22
	(0.67)	(0.54)	(0.47)	(0.62)	(0.60)	(0.71)	(1.90)	(0.76)	(1.17)	(2.07)	(0.94)
I/TS	0.36	0.20	0.13	0.27	0.04	0.02	0.21	0.37	0.09	0.15	0.20
	(0.71)	(0.30)	(0.27)	(0.42)	(0.05)	(0.08)	(0.91)	(1.39)	(0.25)	(0.55)	(0.57)
TL/TS	0.84	0.24	0.75	0.25	0.48	0.30	1.85	2.71	0.99	0.94	0.63
	(1.31)	(0.69)	(1.82)	(0.61)	(0.97)	(0.78)	(11.06)	(9.16)	(3.24)	(2.61)	(3.52)
IP/(IP+EBITDA)	0.24	0.24	0.18	0.24	0.19	0.24	0.22	0.42	0.17	0.23	0.24
	(0.97)	(0.99)	(0.84)	(1.17)	(0.99)	(0.76)	(1.02)	(0.87)	(0.88)	(0.87)	(0.99)
Dividend	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)
Remarks 1	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)
Remarks 2	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	0.15	0.13	0.12	0.08	0.08	0.17	0.09	0.09	0.13	0.12	0.11
	(0.24)	(0.20)	(0.20)	(0.23)	(0.34)	(0.23)	(0.37)	(0.19)	(0.28)	(0.34)	(0.25)
TL/TA	0.70	0.69	0.71	0.74	0.87	0.72	0.67	0.79	0.64	0.65	0.70
	(0.32)	(0.29)	(0.28)	(0.34)	(0.48)	(0.31)	(0.57)	(0.37)	(0.35)	(0.45)	(0.35)
LA/TL	0.46	0.39	0.42	0.40	0.36	0.45	1.33	0.40	0.85	1.04	0.57
	(0.99)	(0.84)	(0.71)	(0.89)	(0.67)	(0.93)	(4.11)	(1.29)	(1.65)	(2.32)	(1.38)
I/TS	0.23	0.14	0.09	0.20	0.03	0.01	0.23	0.21	0.05	0.10	0.12
	(0.47)	(0.22)	(0.22)	(0.31)	(0.05)	(0.05)	(0.94)	(1.09)	(0.18)	(0.42)	(0.39)
TL/TS	0.64	0.47	0.50	0.31	0.45	0.43	4.34	4.80	0.75	0.63	0.68
	(1.15)	(1.69)	(1.45)	(0.83)	(1.00)	(1.41)	(26.1)	(18.4)	(2.88)	(1.98)	(5.64)
IP/(IP+EBITDA)	0.15	0.15	0.12	0.18	0.17	0.14	0.19	0.30	0.10	0.12	0.15
	(0.66)	(0.68)	(0.69)	(0.85)	(0.77)	(0.53)	(0.89)	(0.68)	(0.75)	(0.74)	(0.75)
Dividend	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)
Remarks 1	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)
Remarks 2	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; REMARK1 = 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; REMARK2 = a dummy variable taking the value of 1 if the firm is in various tax arrears; TTLFS = a dummy variable equal to 1 if the firm has not submitted an annual report in the previous year, and 0 otherwise.

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
Micro 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)
Remarks 1	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)
Remarks 2	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)
Divident	-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 R2	0.39	0.29	0.36	0.32	0.31	0.38	0.40	0.33	0.38	0.41	0.34
Pseudo R2 | 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 R2	0.50	0.51	0.47	0.53	0.52	0.38	0.27	0.52	0.45	0.56	0.49
Industry R2 | 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)
Remarks 1	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)
Remarks 2	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)
Divident	-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 R2	0.88	0.95	0.95	0.97	0.85	0.89	0.84	0.86	0.94	0.83	0.96
Industry R2 | 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) **[Skriva in hur McFadden pseudo R2 beräknas]**

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

Model	**RMSE**											
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
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 Tabel 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
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 are computed as one-step-ahead forecasts for the period 2000*Q*1-2002*Q4*. No model is estimated on data after 1999*Q*4. Industry aggregate RMSEs are 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.

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

Industry specific parameters

	Agriculture	Manufacturing	Construction	Retail	Hotel & Restaurant	Transport	Bank. Finance & Insurance	Real-Estate	Consulting & Rental	Not Classified	Industry aggregate
Pseudo R^2	0.40	0.34	0.51	0.38	0.41	0.44	0.43	0.58	0.41	0.46	-
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**

Aggregate model parameters

	Agriculture	Manufacturing	Construction	Retail	Hotel & Restaurant	Transport	Bank. Finance & Insurance	Real-Estate	Consulting & Rental	Not Classified	Aggregate
Pseudo R^2	0.46	0.35	0.51	0.37	0.35	0.45	0.45	0.47	0.44	0.43	0.42
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 are 2,614,248. The parameters used are the ones estimated in Table 3. The decile test numbers in the table are obtained by sorting the estimated default probabilities, in descending order. and computing the default frequencies in the different deciles of the sorted data. Industry aggregate numbers are obtained by generating the estimated default probabilities with the industry specific parameters, adding the observations to a single dataset. and then apply the procedure outlined above to compute the default frequencies for the various deciles.

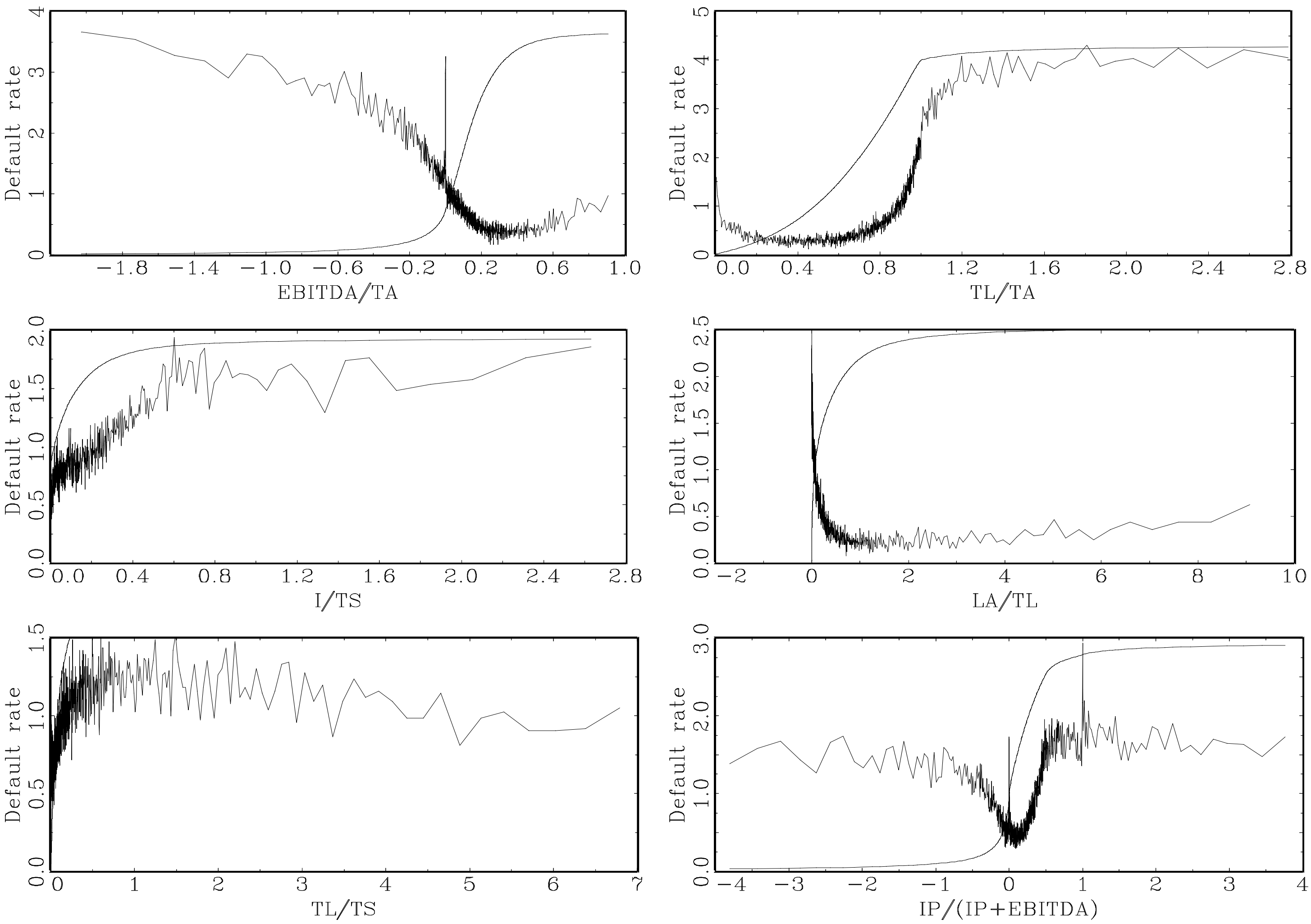

Figure 1: Default rates and the cumulative distribution functions for the accounting data.

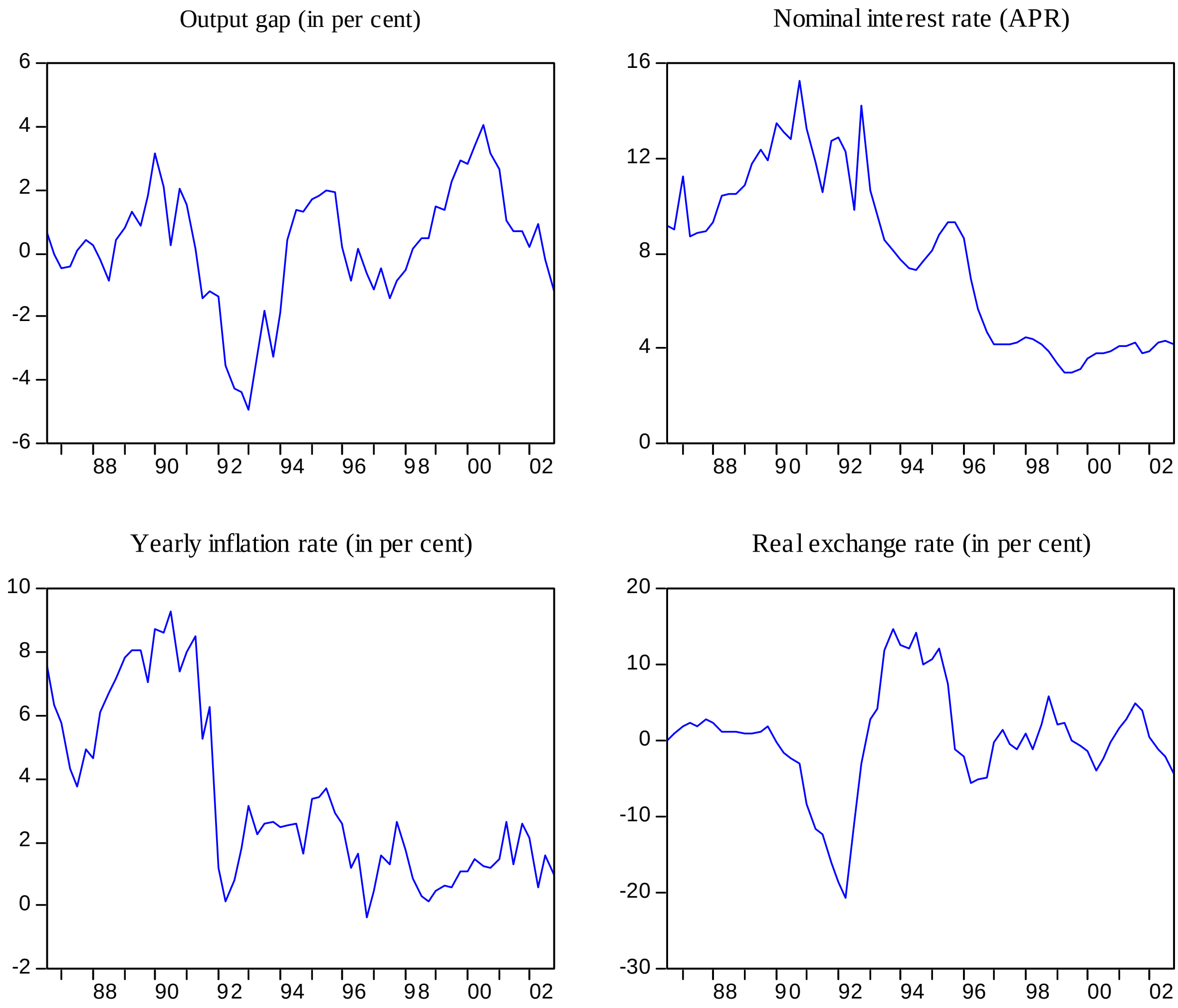

Figure 2: Macro data used in the estimated models.

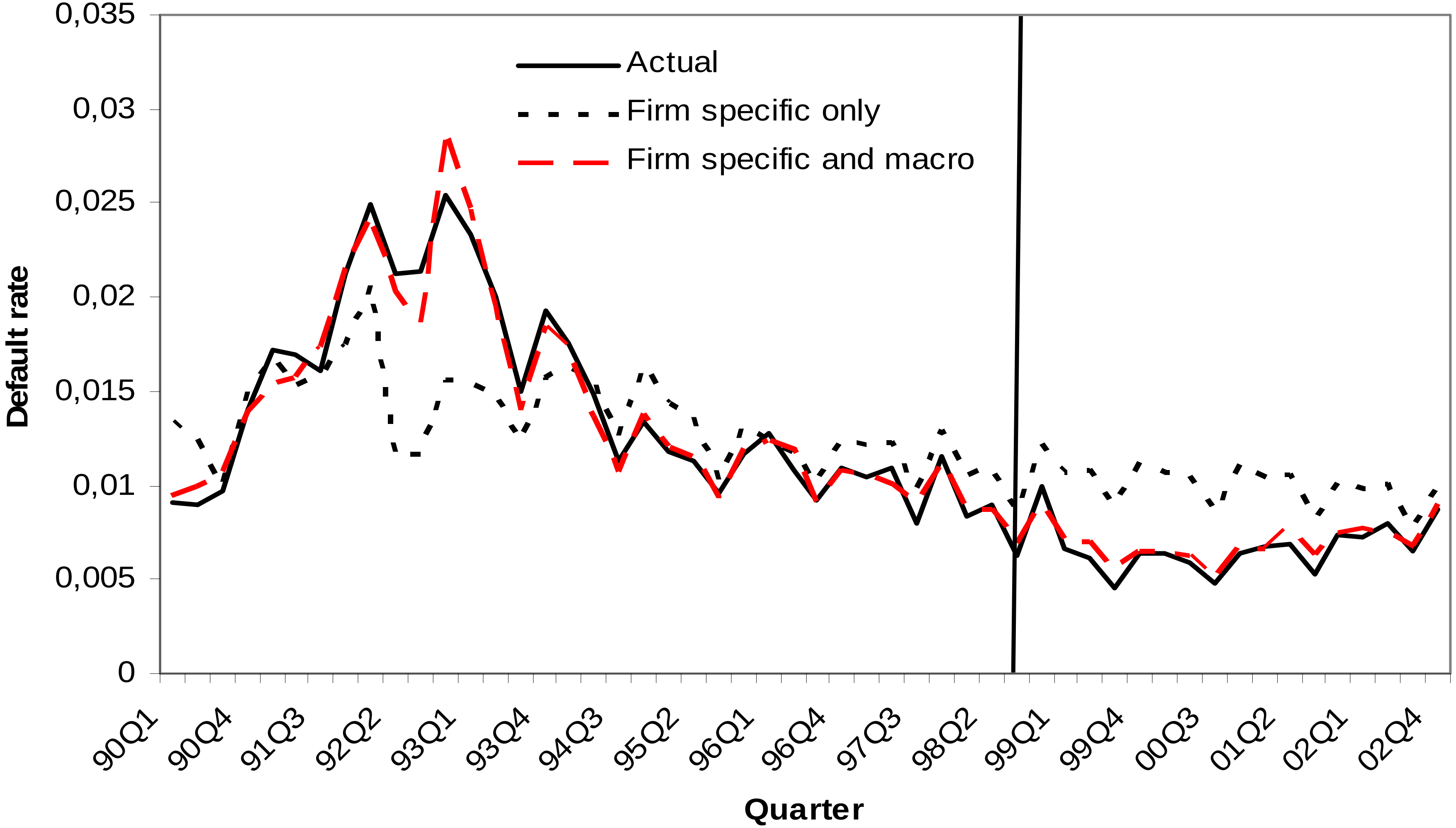

Figure 3: : Actual (solid) and projected (dashed, dotted) aggregate default frequency rates 1990Q1-2002Q4. The projected rates are constructed using the estimated Economy-Wide models in Table 2 (dotted) and Table 3 (dashed). The models are estimated on data until 1999Q4, hence the projections to the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

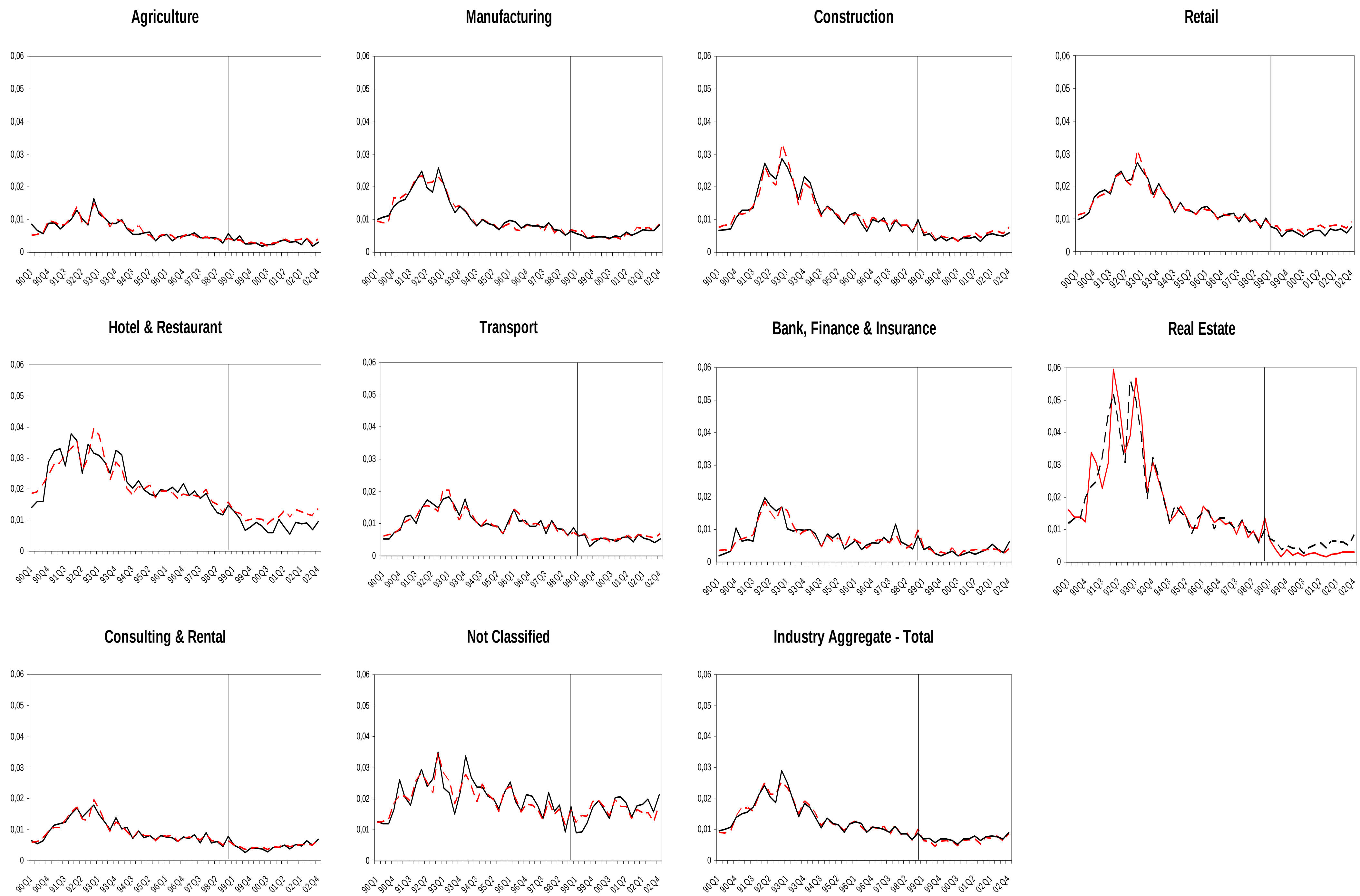

Figure 4: Actual (solid) and projected (dashed) industry default frequency rates 1990Q1-2002Q4. The projected rates are constructed using the estimated industry-specific models in Table 3. The models are estimated on data until 1999Q4, hence the projections to the right of the vertical line are out-of-sample.

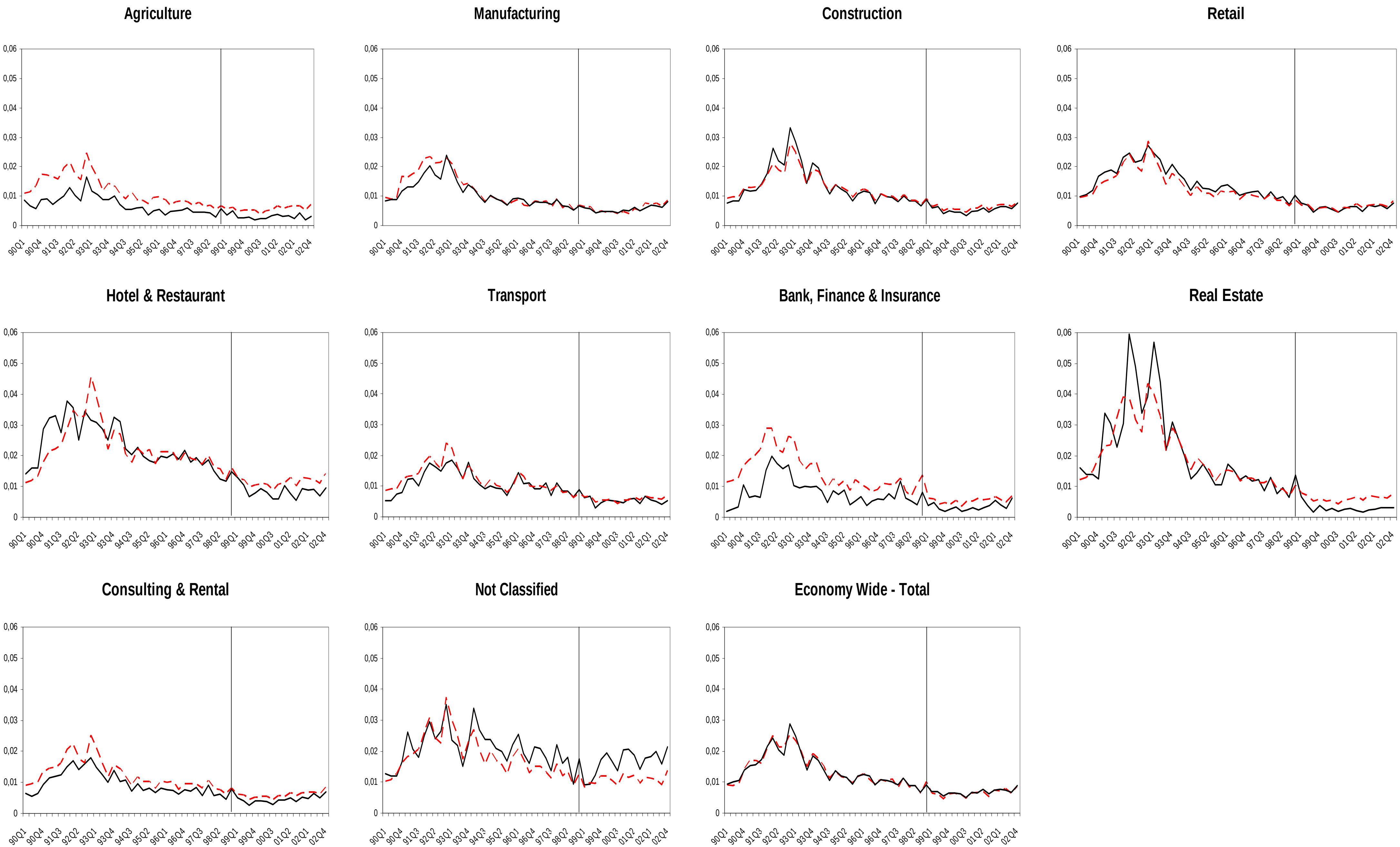

Figure 5: Actual (solid) and projected (dashed) industry default frequency rates 1990Q1-2002Q4. The projected rates are constructed using the estimated Economy-Wide model in Table 3. The model is estimated on data until 1999Q4, hence the projections to the right of the vertical line are out-of-sample.

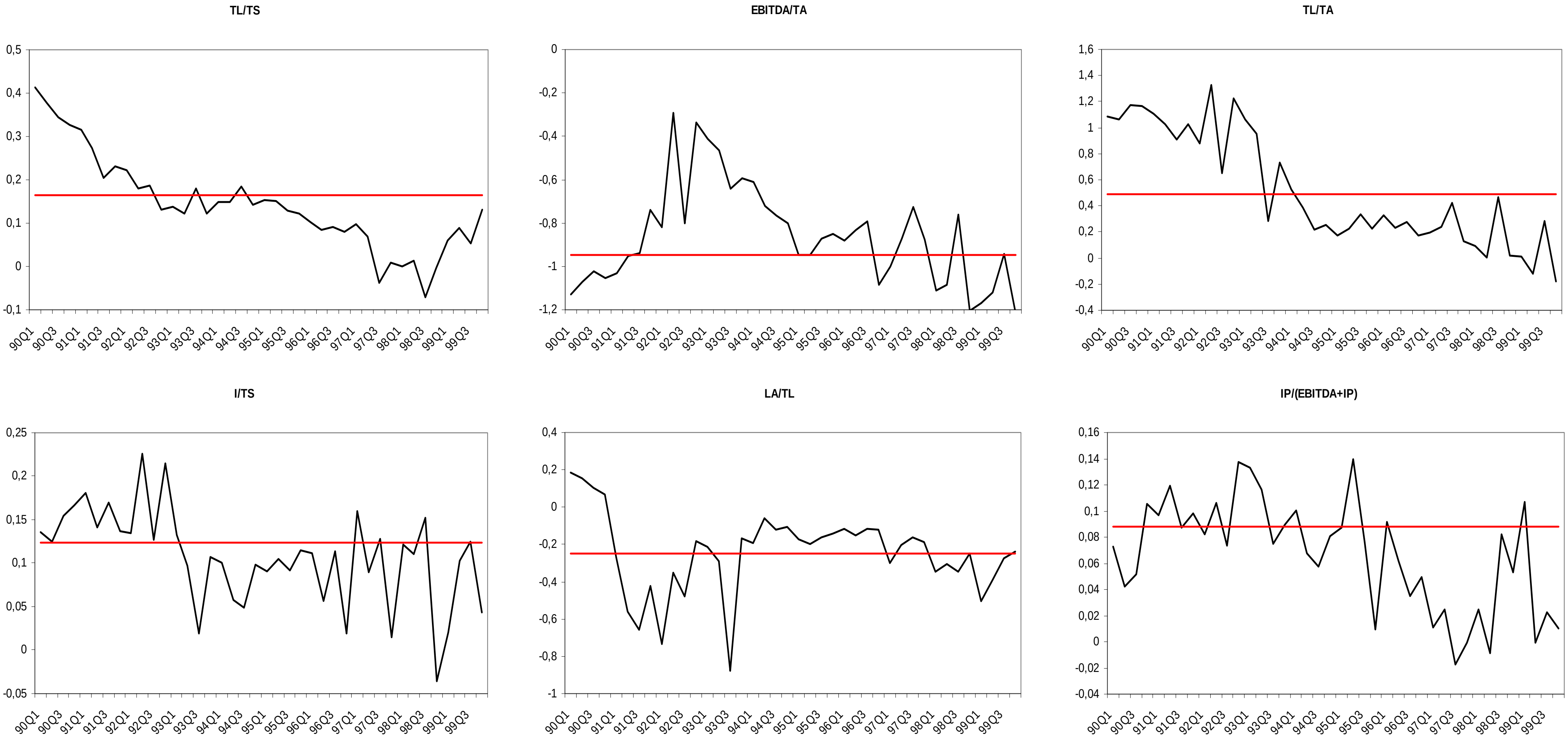

Figure 6: Time-varying coefficients for the accounting ratios in the Economy-Wide model estimated in each quarter without the macro variables included. The red horizontal lines correspond to the estimated coefficients in the Economy-Wide model in Table 2.

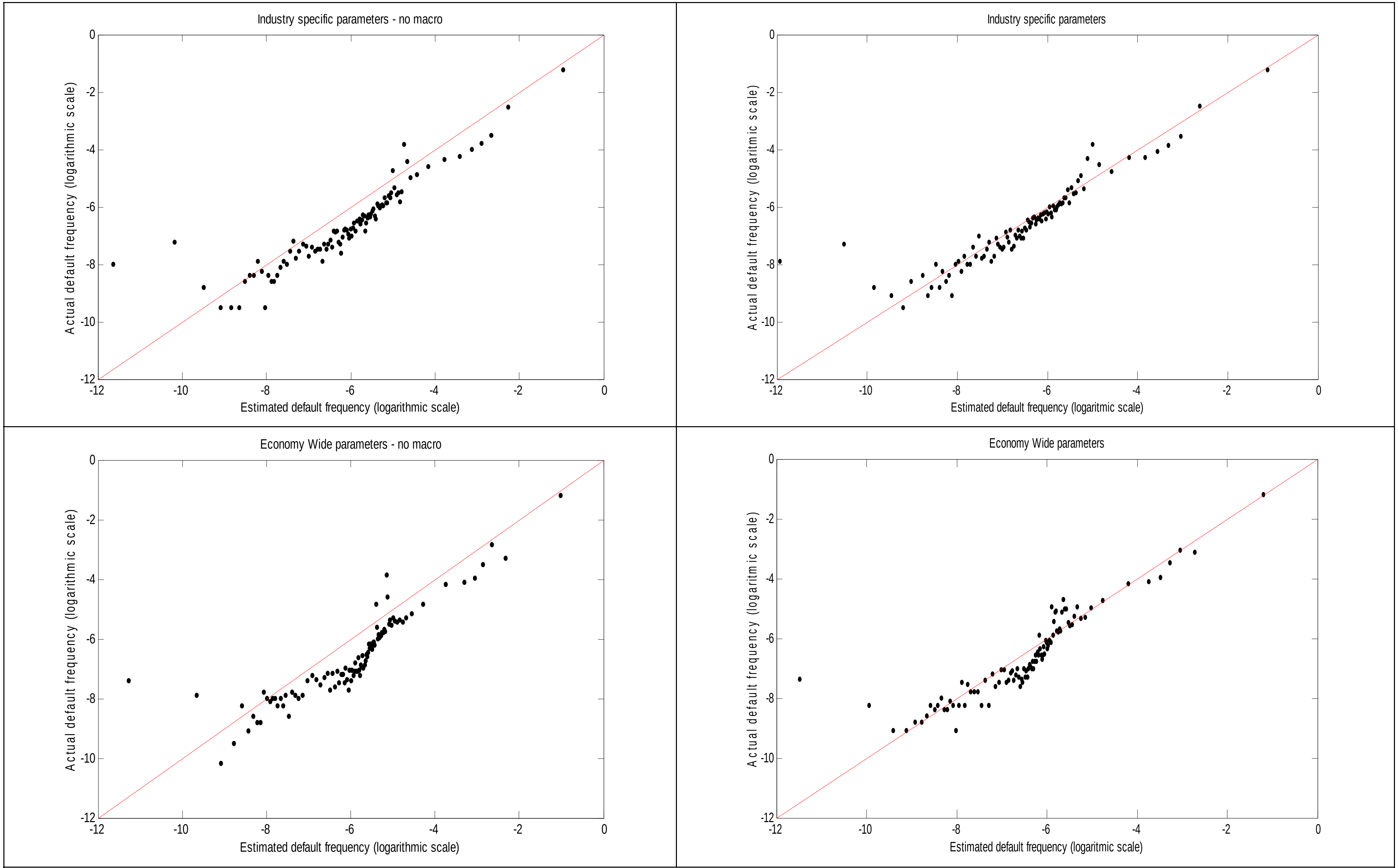

Figure 7: Sorted estimated default percentiles versus actual default frequencies for both economy-wide and industry-specific parameters. Left panel: Only firm-specific variables included (Table 2 models); Right panel: Macrovariables included (Table 3 models).