Macroeconomic Fluctuations and Corporate Financial Fragility*

C. Bruneau,† O. de Bandt,‡ W. El Amri§

This version: July 2009

*The contribution of Meryem Yankol at an early stage of the project is gratefully acknowledged, as well as comments by Valérie Mignon, and seminar participants at Banque de France, Paris School of Economics, the European Meeting of the Econonometric Society, SUERF, the French Economic Association's Annual Conference and the International Symposium in Forecasting Annual Conference. Many thanks to Mireille Bardos for advices during the project and for making a sample from the Fiben database available. All errors remain those of the authors. The opinions expressed are not necessarily those of the Banque de France.

†Banque de France and University of Paris X (*EconomiX*), cbruneau@u-paris10.fr.

‡Banque de France, E-mail olivier.debandt@banque-france.fr, corresponding author : Banque de France 46-1405 DCPM, 75049 Paris Cedex 01 France.

§Banque de France and University of Paris X (*EconomiX*), widad.elamri@banque-france.fr.

Abstract

Using a large sample of accounting data for non financial companies in France on the period 1990-2004, the paper studies the interactions between macroeconomic shocks and companies' financial fragility. We consider links in both directions, namely whether firms' bankruptcies are affected by macroeconomic variables, and whether bankruptcies determine the business cycle. We estimate forecasting equations for firms' bankruptcy using Schumway's (2001) approach and study the joint dynamics of defaults and macroeconomic variables. We provide evidence of significant "second round" effects, with a persistent impact of the output gap on defaults and a feedback effects of defaults on the output gap in a Panel-VAR framework. We illustrate how the model can be used for stress testing.

Key words : financial fragility, macroeconomic shocks, corporate bankruptcies, duration model, stress testing

JEL : G33, E32, D21, C41

1 Introduction

The objective of the paper is to investigate the interaction between macroeconomic cycles and microeconomic shocks, focusing on the firm level. In particular we consider how the financial fragility of firms affects the business cycle, which may itself determine the financial situation of firms. Such a question is in particular relevant in two areas: first in terms of macroeconomic forecasting with a view to incorporating information at the microeconomic level; second, for the implementation of "stress tests" where one considers the response of the financial sector to large macroeconomic shocks. As usually acknowledged, the drawback of the latter approach is that these tests are usually carried out in a static way, omitting the so-called "second round effects" of the shocks to the real economy : a given initial macro-economic shock impacts on the financial situation of firms, which then affects macroeconomic variables. The contribution of the paper is therefore to study the transmission of shocks, and in particular those measuring financial fragility, as defined by the likelihood of corporate defaults.

Here we consider two strands of the literature. First, several papers investigate how financial variables and in particular the financial situation of corporate firms affect the business cycle, in the line of Bernanke and Gertler (1989). Among others, Lown and Morgan (2006) provide evidence that indicators of financial fragility, as measured by business failure rates, together with credit standards have explanatory power for future growth of bank loans and GDP, on top of standard measures of interest rates on loans. Second, there is a growing literature on the impact of macroeconomic variables on corporate defaults (Bordes and Melitz, 1991, Allen and Saunders, 2004, Misiona and Tessier, 2007). Some of them also consider dynamic feedback relationships, as Koopman and Lucas (2003).

Carling et *al.* (2004, 2007) also investigate these issues in the case of the corporate sector in Sweden; they estimate current year or one year ahead default equations for individual firms and measure the effect of the aggregate default probability in a VAR model which also includes output, inflation, the nominal interest rate and the exchange rate. Our study is close in spirit to their approach, but we provide several extensions. We develop a similar analysis in the case of France, using the Banque de France FIBEN database.

In addition we rely on the Shumway's (2001) duration model, which allows to estimate the relationship between macroeconomic variables and defaults over several periods and not only period-by-period. Such models provide more reliable estimates of default probabilities than the usual LOGIT models, since they take into account the progressive deterioration of financial conditions in the corporate sector.

The paper is organized as follows. In section 2 we explain our modelling choices. In section 3 we present the data and the main results we obtain about the bilateral effects of macroeconomic conditions on bankruptcies. Variant scenarios and stress tests are considered in section 4. Section 5 concludes.

2 Modelling choices

In order to study the dynamic impact of financial fragility on the business cycle, Carling et *al.* (2004) use a VAR model where the output gap and other macroeconomic variables are included together with indicators of financial fragility. We follow their approach, but we introduce several differences in the modelling choices as explained hereafter.

The indicators of financial fragility are drawn from the FIBEN database from Banque of France (see section 3) and are available at the yearly frequency over 15 years, while the macroeconomic variables are observed at the quarterly frequency. We estimate a LOGIT type model at the firm level to measure the influence of the business cycle on the default probability (more precisely, on the logarithm of the individual odd ratio).

Due to the relatively short time-span of available data from FIBEN, we decided to focus on the sector level. Thus, we suppose that all sectors are homogenous regarding the determinants of the default rate and accordingly associated with the same LOGIT type model. Sector results are extracted from the firm specific LOGIT model by simple aggregation. Further analysis -not carried out in the paper- could use different specifications of the LOGIT models for the different sectors.

Concerning the reverse impact of financial fragility on the business cycle and more precisely on the output gap, we estimate the parameters of the related linear regression by using PANEL-GMM method. We allow therefore for sector specific dynamics of the business cycle, although its elasticity to the financial fragility indicator is homogeneous across sectors.

In the following section, we describe in detail the system of equations that we use in our analysis.

2.1 The dynamic system

Rather than using a standard VAR model, we refer to the following system of two equations:

$$\begin{bmatrix} \tau_{def,t} \\ X_t \end{bmatrix} = \begin{bmatrix} 0 & A_{12}(L) \\ A_{21}(L) & A_{22}(L) \end{bmatrix} \begin{bmatrix} \tau_{def,t} \\ X_{t-1} \end{bmatrix} + d(L) \begin{bmatrix} Z_t \\ 0 \end{bmatrix} + \begin{bmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{bmatrix}. \quad (1)$$

$\tau_{def,t}$ is an observable indicator of financial fragility contrary to the estimated indicator chosen by Carling et *al.* (2003). More precisely $\tau_{def,t}$ is the empirical counterpart of the logarithm of the odd ratio $\log(\frac{p_{def,t}}{1-p_{def,t}})$,[1] derived from the observed default frequency.[2] $A_{12}(L)$, A_{21} (L) and $A_{22}(L)$ are lag-polynomials.

[1] It is usual to define the odd ratio as :

$$\frac{p_{def,t}}{1-p_{def,t}},$$

and to consider the logarithm of this ratio, which is an increasing mononotic transformation of the default probability $p_{def,t}$.

[2] We should define $\tau_{def,t}$ as $\log(\frac{f_t}{1-f_t})$ with f_t denoting the default frequency at year t but we keep the usual notation $\log(\frac{p_{def,t}}{1-p_{def,t}})$.

X_t are macroeconomic variables and Z_t is an aggregate counterpart of individual financial information Z_{it} measured for each firm i at date t.

The two equations are estimated separately. We take into account the contemporary correlation between ϵ_{1t} and ϵ_{2t} by introducing $\tau_{def,t}$ into the second equation ($A_{21}(0) \neq 0$) but there is no contemporaneous variable in the first equation.

In order to compensate for the short time-span of available data (15 annual observations were made available to us and 11 could effectively be used), we use sector-based series to estimate the second equation. Accordingly, it should be written as:

$$X_{jt} = A_{21}(L)\tau_{def,jt} + A_{22}(L)X_{jt-1} + \epsilon_{2,jt},$$

where j is a sector index. It is estimated by Panel-GMM as indicated in section 3.2. As outlined before, we allow for heterogeneity across sectors in the estimation.

The first equation:

$$\tau_{def,t} = A_{12}(L)X_{t-1} + d(L)Z_t + \epsilon_{1t},$$

is estimated from a LOGIT-type model specified at the individual (firm) level:

$$\tau_{def,it} = Log\frac{p_{it}}{1-p_{it}} = c + A_{12}(L)X_{t-1} + d(L)Z_{it} + \sum_{j=1}^{J-1} c_j \mathbf{1}_j + \epsilon_{1it},$$

where p_{it} denote the default probability for firm i at date t and $\mathbf{1}_j$ is equal to 1 if firm i belongs to sector j and 0 otherwise. The c_j coefficients are associated with the sector fixed effects.

Then the estimates of the parameters are used at the sector level, under the assumption that each sector is homogenous :

$$\tau_{def,jt} = \sum_{i \in \mathrm{sector}_j} Log\frac{p_{it}}{1-p_{it}} = c + c_j + A_{12}(L)X_{t-1} + d(L)\sum_{i \in \mathrm{sector}_j} Z_{it} + \epsilon_{1,jt}.$$

In the previous regression, we do not introduce any sector-based variables, apart from the fixed effects; cyclical business fluctuations are captured through aggregate macroeconomic series X_t. The individual dimension is taken into account through the Z_{it} which include various financial indicators.

Notice that we use interchangeably bankruptcy and defaults, defined as the termination of business, which is different from payment incidents on commercial debt (as opposed to incidents on financial debt leading to the termination of business after a legal procedure). We use incidents on commercial debt as leading indicator of default, as mentioned below. Note also that individual financial information Z_{it} is taken as exogenous. A further step, not carried out in the paper, would imply either to assess the impact of macroeconomic variables X_t on Z_t and also including a richer set of sector specific variables in the individual LOGIT, or run sector specific LOGIT models.

As indicated above, to estimate $A_{12}(L)$ and $d(L)$, we refer to a LOGIT–type approach. However, as made clearer below, it is worth emphasizing that this approach is different from the usual LOGIT one.

Indeed, in order to capture the macroeconomic information content for predicting future defaults, it is decisive to investigate the bankruptcy events in a dynamic framework, which is not allowed in the standard LOGIT framework, where a defaulting company only appears once in the likelihood.

The LOGIT-type model we refer to is a particular duration model introduced by Shumway (2001), whose main principles are exposed in the following section.

2.2 Duration models and the multiperiod LOGIT model of Shumway (2001)

To explain the default risk, one often refers to a latent variable which is the ability of the company to satisfy its financial debt obligations. If the latent variable is smaller than a critical value (which can be supposed to be equal to 0), the firm defaults ($y_i = 1$, and $y_i = 0$ otherwise). A duration model can provide a dynamic characterization of these default events, as it measures the life duration of each firm - i. e. the number of years before bankruptcy- as a function of relevant micro or macro variables, which can be observed during the whole life of the firm. In particular, such a model is able to capture the progressive worsening of the financial conditions.

Moreover the duration model, once estimated, can provide forecasts of default at different horizons. This can be useful if one wants to define different indicators of financial fragility for example if one looks at default events over a period instead of default at a given point in time.

In what follows, we will estimate the probability of defaulting at year t, when the company is still in operation at year $t-1$, on the basis of information available two or three years before. This choice is essentially justified by data quality concerns, since the reliability of accounting data for companies close to bankruptcy or the year of default may be questionable. Indeed, either the company does not provide any accounting information, or this information is not reliable, or the company has already been restructured and is a different entity in comparison to the original company.

More precisely, the duration model we use to forecast defaults, is specified and estimated as follows.

One observes n firms over the period $[0, T]$ and for each firm i, one defines the life duration, D_i, that is the duration before default. If t_i denotes the default date and $\underline{t_i}$ the date of the start of its activity, one has:

$$t_i = \underline{t_i} + d_i,$$

where d_i is a realization of the random variable D_i. The date at which firm i leaves the sample without defaulting is noted c_i. To exclude "left side"censored data, one can suppose that $\underline{t_i} \geq 0$. Note than one can have right side censored data when firms leave the sample at a date c_i for a cause which is not default (for example, a merger).

Let us assume that the population is homogeneous and that the default events are independent across different firms. Such an assumption is a priori questionable but becomes acceptable provided that the default events are investigated "conditionally on macroeconomic information", which corresponds to our specification choice.

Let f denote the density distribution function of D_i; it is identical for all firms, according to the homogeneity hypothesis. A firm i contributes to the likelihood by $f(d_i)^{y_i}$ (*i. e.* $y_i = 1$) if it defaults at $t_i = \underline{t_i} + d_i \leq T$. Otherwise, it contributes by $S(c_i - \underline{t_i})^{1-y_i}$, where S denotes the survival function[3].

Accordingly, the likelihood can be written as:

$$\begin{aligned} L(u_1..., u_n) &= \prod_{i=1}^{n} f(u_i - \underline{t_i})^{y_i} S(u_i - \underline{t_i})^{1-y_i}, \\ \text{where } u_i &= Min(t_i, c_i). \end{aligned}$$

If the firms leave the sample only because of default events, one has $\{u_i = t_i < T;\ y_i = 1\}$. But $\{u_i = c_i = T;\ y_i = 0\}$ if firm i does not default over the sample period. If firms can leave the sample for other reasons, one can have: $\{u_i = c_i < T\ ;\ y_i = 0\}$.

When the duration variable is discrete ($d^{(k)}$ takes integer values for $k = 1, 2, ...$), one can prove that the likelihood can be written only by introducing the hazard function h[4]:

$$L(u_1, ..., u_n) = \prod_{i=1}^{n} h(u_i - \underline{t_i})^{y_i} \prod_{d^{(k)} < u_i - \underline{t_i}} (1 - h(d^{(k)})).$$

It is possible to specify the hazard as a function of observable variables which can be macroeconomic variables or variables measured at the firm level. In what follows, we choose the specification:

$$h_i(d, Z_{i,\underline{t_i}+d-H}, X_{\underline{t_i}+d-H'}) = \frac{1}{1 + \exp(Z'_{i,\underline{t_i}+d-H}\beta_1 + X'_{\underline{t_i}+d-H'}\beta_2 + c)}.$$

where X denotes a set of K macroeconomic variables with K strictly smaller than T. H and H$'$ are the lags introduced for, respectively, the microeconomic and the macroeconomic explanatory variables. These variables are included with lags in order to account for delays in the availability of the related information. In the literature, H is often defined as the horizon of the projection : default at t is explained by financial information at $t - H$. Eventually, we get an estimate of the default probability that we have defined as $p_{it} = 1 - h_{it}$.

Now, if one defines the observations or "individuals" as year-firms and no longer firms, one can prove that the likelihood of the duration model has the same form as the likelihood

[3] The survival function is defined as:

$$S(d) = P(D > d) = 1 - F(d),$$

where F is the cumulated distribution function of the D variable.

[4]

$$h(d) = \frac{f(d)}{1 - F(d)} = \frac{f(d)}{S(d)} \Leftrightarrow f(d) = h(d)S(d).$$

function written for the so-called multiperiod LOGIT model examined by Shumway (2001) with the F_{LOGIT} function specified as:

$$F_{LOGIT}(z'_{(i,t)}\beta) = 1 - h(z'_{(i,t)}\beta),$$

where the dependent variable $y_{(i,t)}$ is equal to 1 if firm i defaults at time t and 0 otherwise; the $z_{(i,t)}$'s account for all the explanatory variables. See Annex A for details.

With the chosen specification for h, the parameters of the duration model can be estimated by using a standard LOGIT procedure.

However, it is worth noting that the individuals are supposed to be independent in such a procedure. It is of course not the case, when one considers year-firms as individuals. Indeed, observations related to the same firm but to different dates are not independent. This is the reason why Shumway (2001) suggests to correct the number of degrees of freedom for chi-square distributions used in the standard tests.[5]

In the following section, we present the data and the main estimation results. Details on the method used to estimate the second equation (GMM panel estimation in the lines of Blundell and Bond, 1998) are reported in Annex B.

3 Data and estimation results

We now present the data that we use, and summarize successively the estimation results obtained for the two equations of our system.

3.1 Data sources

Two types of data are used, financial information on firms as well as data on bankruptcies. Regarding individual financial data on non financial companies in France, several sources are available, either from Insee or from the Banque de France FIBEN database. There are also different sources on corporate bankruptcies, either from Insee or Banque de France. We use an unbalanced sample of individual companies from the FIBEN database, for which we have information on the date and cause of exit from the sample. This allows to measure precisely the occurrence of defaults. We check that our sample is representative of national developments using the comprehensive data published by Banque de France on corporate defaults at the sector level, concentrating on the 1990-2006 period.[6] We also compare our data to those published by Insee, referring to the study by Dommens (2006). The sector data from Banque de France are quite consistent with those from Insee at the sector level, although Banque de France also uses specific information. In order to compute sector default rates, the aggregate number of failures by sector is divided by the number of companies by sector, using Insee data from the Alisse database.[7] We adjust

[5] The right number of the degrees of freedom is the number of firms (provided that the likelihood is written conditionally on the macroeconomic factors, which justifies the independence assumption for different firms).

[6] See in particular data on bankruptcies published by Banque de France at the sectoral level: http://www.banque-france.fr/fr/stat_conjoncture/telechar/publi/defaill_entrep.pdf

[7] The data come from the ALISSE database available from Insee at

the level of the computed default probabilities to ensure that they are identical with the ones published by Nahmias (2005) for the years 2002 to 2004. It should be kept in mind that the FIBEN database excludes very small companies and is therefore less complete than the set of income tax returns used, e.g., by Crépon (1993). Nevertheless the full set of FIBEN data is referred to as "Fiben exhaustive" in Table 1 below.

Our sample of individual data from the FIBEN database initially included yearly accounting information on a sample of 259,890 non financial companies in France on the 1990-2004 period (hence a total of 1 551,003 accounting statements) with a total of 35,875 defaults. These data are referred to as "Fiben sample" in the Table 1 below.

As it is typical with individual data, the database was filtered for outliers since there are a number of extreme values among the observations of the financial ratios constructed from raw data. To ensure that statistical results are not heavily influenced by outliers, we replace all observations with value above the 99th percentile of each variable by that value. All values lower than the first percentile of each variable are "winsorised" in the same manner.

Financial accounts are published continuously over the year. Most of them cover the preceding year until 31 December, but may be published with delay. In addition some companies close their account during the year. We assume therefore that financial accounts published until 30th June of year N, actually cover operations for year $N-1$.[8]

When comparing the bankruptcy data in our sample with the ones available at the sector level, it appears that only data on bankruptcies since 1994 are reliable, hence a total of 11 years. We decide therefore to concentrate on the post-1994 period.

We also use data on payment incidents or defaults on commercial debt, collected by Banque de France. Such events do not automatically lead to a bankruptcy (as measured by default on financial debt), but can be viewed as an harbinger for future bankruptcy.

3.2 Impact of bankruptcies on the business cycle

Given the small time dimension, we use the sector dimension to increase degrees of freedom. To measure the business cycle, we use data on value added at the sector level from the Insee National Accounts to construct an output gap indicator.[9] We estimate one dynamic panel equation where the output gap (GAP), is explained by its lagged value as well as the observed τ_{def}. The equation is estimated by GMM, using Blundell and Bond's (1998) System GMM method (see Annex B), since OLS is biased given the correlation between the error term and the lagged endogenous variable.

We estimate the second equation of system (1), by regressing the fragility indicator τ_{def} as computed from the three databases, respectively from Insee, the exhaustive Fiben

http://www.alisse.insee.fr/SelectionFD.jsp?item=SERIES

[8] More precisely, data available between 1*st* January of year N and 30*th* June of year N will appear in Year $N-1$.

[9] The output gap at the sector level is computed as deviation from a linear trend on the logarithm of sector value added, using data from Insee National Accounts (working day adjusted) at the sector level.

database and the "FIBEN sample[10]" that we have used to estimate our LOGIT type model. The results are quite similar.

The use of three sources of default rates in the estimation of the second equation allows us to compare our results to those from more complete databases.

Table 1 : Dynamic Panel estimates of the output gap at the sector level (GAP_j)

	Insee		*Fiben* (*exhaustive*)		*Fiben* (*sample*)	
Variable	*coef.*	*s.e.*	*coef.*	*s.e.*	*coef.*	*s.e.*
$GAP_{j,-1}$	1.0804***	0.0670	1.0569***	0.0649	0.8574***	0.0909
$GAP_{j,-2}$	−0.5198***	0.0521	−0.5032***	0.0548	−0.3222***	0.0835
τ_{def}	−0.0008**	0.0003	−0.0010*	0.0920	−0.0023*	0.0013
m_1	0.02		0.02		0.04	
m_2	0.11		0.12		0.41	
Sargan$^{(1)}$	0.37		0.36		0.126	
Nb. obs.	120		120		100	

Notes : ***,** and * denote significant at 1, 5 and 10% levels, respectively.$^{(1)}$ $p-value$ of Sargan's overidentification test of instruments; Estimation : Dynamic panel estimation with Blundell and Bond (1998)'s 1-Step method using System GMM estimator; Period: 1995-2006; m_1 and m_2 are $p-values$ of Arellano-Bond tests of respectively $AR(1)$ and $AR(2)$ serial correlation; $\tau_{def} = \log(p_{def} \;/1 - p_{def})$.

As indicated in Table 1, when considering an AR(2) model for the output gap, the sector default ratio appears to contain additional information.[11]

We find that the signs of the coefficients are stable and their values are relatively close for the different cases. We do not find any serial correlations of order 2 (Arellano and Bond test). Moreover the instruments are validated by the of Sargan /Hansen test.

3.3 Evidence of the Impact of the business cycle on bankruptcies from a multiperiod Logit model

We now concentrate on default in year t given that the company has not defaulted at $t-1$, conditional on financial information available at $t-3$ and macroeconomic information available at $t-2$ or $t-3$. We therefore estimate a multiperiod LOGIT model with $y_{(i,t)} = 1$

[10] This is the database used by Banque de France for the calculation of scores for the financial assessment of companies.

[11] When considering additional macroeconomic variables (long term interest rate and inflation) defaults remain significant and the coefficient is quite similar (results are available upon request). However the use of a more comprehensive model of GAP for stress testing is left for future work. In section 4, we only consider shocks to inflation and interest rate channeled to defaults through the first equation (see system (3) below).

if firm i defaults at t (with i still operating at date $t-1$) and with lagged independent variables $X_{t-2(3)}$, for macroeconomic variables and $Z_{i,t-3}$ for financial information.

Accordingly, the default probability underlying the indicator of financial fragility τ_{def} is exactly the hazard function of the duration model.

3.3.1 Explanatory variables

In the specification of the hazard function, as indicated in section 2.2, we have introduced variables with a lag of H periods. In what follows, we concentrate on the case with H equal to 3 for the financial variables because of the poor quality of the accounting data when the company is about to go bankrupt. We consider also shorter horizons.[12] However, the macroeconomic variables are introduced with shorter lags, *i. e.* with lags from 2 to 3.

Indeed, several recent research on defaults stress the importance of long horizon, as Campbell et *al.* (2008) on the US (horizons from 24 to 36 months) as well as Löffler and Maurer (2008), who consider a range of horizons from 2 to 5 years. Other studies include lagged macroeconomic variables and show that they contain significant marginal information, as Bonfim (2007) for Portuguese firms or Duffie et *al.* (2007) for US data.

Concerning the explanatory variables, we investigate several sets of such variables. In particular, the forecasting models incorporate Altman's (1968) explanatory variables, as well as some variables drawn from reports of the Observatoire des entreprises of Banque de France (2004).[13] But contrary to Shumway (2001), we do not introduce any market-driven explanatory variables because we include in the sample a large number of small and medium size companies which do not have access to financial markets.

For each specification of the endogenous variable, we estimate three models: a first one with only financial ratios as independent variables; a second one, including information about pre-default events and debt with two dummy variables:

- IP = 1 if one non-payment incident on commercial debt occurs during the year (= 0, otherwise);
- Tax Arrears =1 if the company is indebted vis-à-vis the government (Tax and Social Security) (= 0 otherwise),

In the third specification, we add macroeconomic variables as explanatory variables, introducing an output gap variable (GAP), the long term interest rate, the inflation rate

[12] As indicated in 3.1, financial information at t-3 corresponds to information published between 1st July of year t-3 and 30th June of year t-2. This is the choice usually made by practionners in France (Banque de France, 2004). Taking financial information at $t-1$ is possible but would imply relying on a smaller data set, due to the non publication of financial information by many companies in the year where the bankruptcy is confirmed by the judge, or the year before, so that these companies have to be excluded from the analysis. Annex E presents results with financial information at $t-2$. Robusteness checks based on t-1 information are available from the authors upon request.

[13] The variables suggested by Altman to predict companies' default are described extensively in Altman (1993). They include the following ratios: working capital to total assets (WC/TA); retained earnings to TA (RE/TA); earnings before interest and taxes to TA (EBIT/TA); market equity to total liabilities (ME/TL); sales to TA (S/TA).

and the dollar/euro nominal exchange rate (amount of euros per 1 USD). These variables are included at year $t-2$ and $t-3$, but impact defaults at year t.

Such an approach allows to measure the information content of the additional variables.

We now provide more details about the definition of the different financial variables that we introduce in the first step model. The ratios that we have retained are the following ones (Annex G contains summary statistics of these variables):

- Profit = Profit / Total assets;
- Leverage = Total Borrowing/Total Liabilities;
- Liquidity = Liquid Assets/ Financial Debt;
- Business credit received (in days of purchases) = 360 * intercompany loans/purchases (VAT included);
- Int = Interest paid/ (Interest paid +Profit).

As mentioned before, we have 1 $551,003$ year-firms and a $35,875$ defaults among them. As we focus on a defaults at 3 years horizon, our sample is reduced to $863,005$**,** year-firms with $13,377$ defaults. The difference corresponds to the removal of companies that default within the first two years of their existence.

3.3.2 Empirical results

Table 2 summarizes the main results for the three year-horizon, obtained for the 1994-2004 period, with financial variables at $t-3$ and macroeconomic variables at $t-2$ and $t-3$ (results for two year-horizon, i. e. to explain defaults at t with financial information at $t-2$ and macroeconomic variables at $t-2$ and $t-1$, are exhibited in Annex E; they are very similar to Table 2).

First, the microeconomic or financial ratios have a significant information content, with the expected sign for the associated β parameter. The dummy variables also have, as expected, a significant (positive) impact on the default probability. The macroeconomic variables have a significant coefficient with the expected sign. GAP enters with a negative sign indicating that lower GAP is associated with higher failures as displayed in Figure 1. An increase in the interest rate implies higher borrowing costs, hence leads to an increase in the likelihood of defaults. A depreciation of the exchange rate vis-a-vis the US dollar, conditional on inflation and the business cycle, signals weaker productive performance and anticipates on more defaults. Moreover, the discriminant power of the model,[14] as measured by the AUC criterion, increases when additional variables are included.

[14] That is its ability to discriminate between the defaulting and not defaulting firms.

Table 2 : Logit models estimation of corporate defaults at date t . Dynamic approach with information at $t-3$

Variable	*Model I* (*Dyn*) coef.	s.e.	*Model II* (*Dyn*) coef.	s.e.	*Model III* (*Dyn*) coef	s.e.
Micro variables						
$Cst.$	-6.0702^{***}	0.0592	-6.3487^{***}	0.0690	-17.7689^{***}	0.3337
$Profit_{t-3}$	-2.7730^{***}	0.0934	-2.6064^{***}	0.0935	-2.7160^{***}	0.0942
$Leverage_{t-3}$	3.2956^{***}	0.0457	3.0779^{***}	0.0464	3.0446^{***}	0.0467
Liq_{t-3}	-0.0091^{***}	0.0009	-0.0082^{***}	0.0009	-0.0069^{***}	0.0009
$Bus.\ Cred._{t-3}$	0.00081^{**}	0.00019	0.00082^{**}	0.00019	0.0006	0.0002
Int_{t-3}	0.2464^{***}	0.0193	0.2230^{***}	0.0193	0.1658^{***}	0.0191
$Size\ (large)_{t-3}$	-1.8121^{***}	0.1000	-1.7036^{***}	0.1001	-1.5553^{***}	0.1003
$Size\ (medium)_{t-3}$	-0.3564^{***}	0.0404	-0.3182^{***}	0.0406	-0.1749^{**}	0.0410
IP_{t-3}			1.1403^{***}	0.0291	1.0774^{***}	0.0297
$Tax\ Arrears_{t-3}$			0.3298^{***}	0.0418	0.3234^{***}	0.0419
Macro variables						
GAP_{t-3}					-0.2472^{***}	0.0208
INF_{t-3}					-0.4338^{***}	0.0287
IRL_{t-3}					0.5265^{***}	0.0327
$Ex.Rate_{t-3}$					0.0543^{***}	0.0026
GAP_{t-2}					-0.3954^{***}	0.0229
INF_{t-2}					-0.5933^{***}	0.0260
IRL_{t-2}					0.1211^{**}	0.0214
$Ex.Rate_{t-2}$					0.0397^{***}	0.0018
$\widetilde{R}^2$ [15]	0.0914		0.1015		0.131	
$LR\ test^{(1)}$			235.86^{***}		998.21^{***}	
$AUC^{(2)}$	0.764		0.773		0.803	
Nb. obs.	$863,005$		$863,005$		$863,005$	

Notes : ***,** and * denote significant at 10, 5 and 1% levels, respectively.; $^{(1)}$Likelihood Ratio test statistic for nested models $\sim \chi^2_{(df)}$; $^{(2)}$Area Under the Receiver Operating Characteristics (ROC) Curve; with $df = 2$ (*resp.* 8) for Model I *vs* II (*resp.* Model II *vs* III)

[15] Nagelkerke's $\widetilde{R}^2$ is obtained as: $\frac{R^2}{R^2_{\max}}$ with $R^2 = 1 - \left\{LnL_{\max}(\beta=0)/\,LnL_{\max}(\widehat{\beta})\right\}^{\frac{2}{N}}$ and $R^2_{\max} = 1 - \{LnL_{\max}(\beta=0)\}^{\frac{2}{N}}$. $\widetilde{R}^2$ measure the discrepancy between the LOGIT models with $(\widehat{\beta})$ or without $(\beta=0)$ the dependent variables X. N is a sample size.

It is usually observed that adding macroeconomic variables in the form of time series may introduce multicollinearity in the model (Beck et *al.*, 1998), and consequently biases on standard errors. We assess therefore the usefulness of the additional variables using a likelihood ratio (LR) test for nested models (under H_0, all coefficients on macroeconomic variables are equal to zero). The null hypothesis of absence of additional information is clearly rejected, as indicated in Table 2. Furthermore, the coefficients on microeconomic variables are stable across models (I, II or III).

Second, based on these models it is therefore possible to compute the aggregate default rate for each year. As Figure 1 shows, the models perform quite well, and the estimated default with model III (black, dashed line) is actually quite close to the observed default (red line). The Figure also displays the output gap (GAP) with an inverted scale. We can conclude from the Figure 1 that defaults estimated with model III, as well as actual defaults are indeed negatively correlated with the output gap. Negative output gaps are associated with higher defaults. Note that such a correlation is not directly the consequence of the integration of macroeconomic variable in the logistic equation, since these variables are introduced in the model with a lag equal to 2-3 periods. For example expected defaults for the year 2000 are based on financial information in 1997 and macroeconomic situation in 1998-1997. However we show that at the aggregate level expected default in 2000 is negatively correlated with the output gap in 2000.

3.3.3 Robustness checks

In Annex F, we test the stability, as well as in and out-of-sample performance, of the optimal model. The estimation of model III on different subperiods provides evidence that the coefficients are stable and that the results are robust.[16] Out-of sample results are quite satisfactory for horizon H= 2 (i.e. default at t with data up to $t-2$ or $t-3$) and better than for H=3.

The use of multiperiod Logit models with observed time series may provide forecasts that are very sensitive to abrupt changes in macroeconomic variables and therefore exhibiting a jagged behavior. To avoid this problem, several studies (Beck et *al.,* 1997, 1998) advocate the use of spline functions to smooth the macroeconomics series and get more realistic out-of-sample forecasts. Eubank (1988) suggests in particular the use of «natural cubic splines». We implement this method.

The estimation of the model with spline transformations of the variables for horizon H=2 and H=3 improves upon our previous results (available upon request from the authors) and offers better in and out-of-sample forecasts for model III. The improvement is mainly visible for H=3 (see Figure 10 of Annex F).

[16] As many other empirical studies, we splitted our sample in two parts corresponding to 2/3 (for estimation) and 1/3 (for out-of-sample analysis) in order not to create an important bias with respect to the original model.

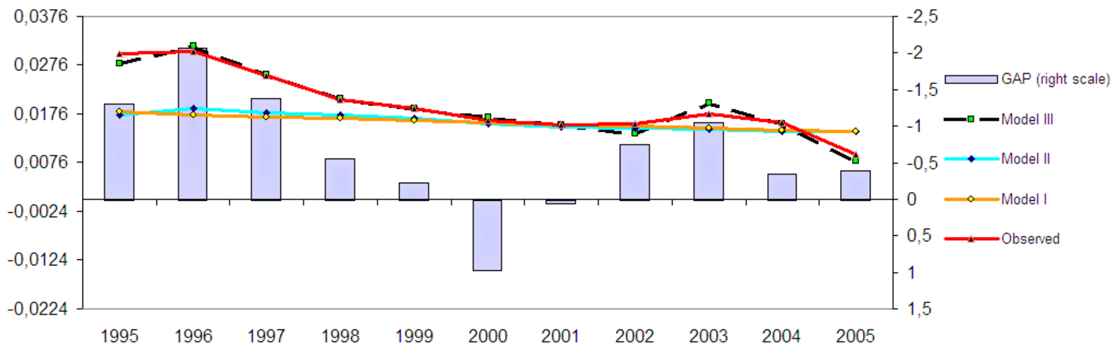

Figure 1 : Aggregate default rate- observed and estimated at t from information at $t-3$ (financial and Macro variables) and $t-2$ (macro economic variables)

4 Variant scenarios and stress testing

Up to now, we have investigated separately the impact of macroeconomic conditions on bankruptcies and the effects of bankruptcies on business cycle. In what follows we aim at examining more in depth both types of links.

We exhibit impulse response functions of GAP and defaults to an innovation to the output gap, that can be conveniently used for running stress tests in a dynamic setting (including second round effects). We also use our duration-Multiperiod Logit model to investigate how shocks to the output gap affect the distribution of defaults.

4.1 Impulse response analysis

For that purpose, we use the system (1) of two equations with defaults (measured as the logarithm of the odd ratio τ_{def} derived from our LOGIT equation on defaults) and output gap (as an indicator of business cycle) at the sector level. As before, we use the sector level in order to increase the number of degrees of freedom, as defaults are only available at the annual frequency. At this stage, we treat our system as an "exogenous VAR", or VAR-X, where the other macroeconomic variables are exogenous to the output gap and the default ratio (this assumption will be partly dropped for stress testing exercises, taking into account the effect of output gap shocks on inflation and interest rates).

One equation, derived from our default Multiperiod LOGIT equation in the previous section, allows to measure the effect of activity on defaults. The second equation measures the impact of defaults on future output gap. Linking the two equations, one can examine how shocks to macroeconomic variables affect defaults, which then impact on macroeconomic variables according to a standard impulse response analysis. This is similar to the approach of Carling et *al.* (2003) and De Graeve et *al.* (2008) who introduce the (estimated) aggregate default rate in a VAR at the macroeconomic level. Here we extend their approach in the panel dimension, on the basis of data on output gap and observed defaults at the sector level.

To summarize the system we use is as follows:

$$\left[\begin{array}{c}\tau^j_{def,t}\\ GAP^j_t\end{array}\right]=\left[\begin{array}{cc}0 & A_{12}(L)\\ A_{21}(L) & A_{22}(L)\end{array}\right]\left[\begin{array}{c}\tau^j_{def,t}\\ GAP^j_{t-1}\end{array}\right]+d(L)\left[\begin{array}{c}Z^j_t\\ 0\end{array}\right]+\left[\begin{array}{c}\epsilon_{\tau^j_{def,t}}\\ \epsilon_{GAP^j_t}\end{array}\right], \quad (2)$$

where $\tau^j_{def,t}$ denotes the indicator of financial fragility for sector j at date t defined from the default frequency observed at the sector level, as explained before.

Our system is estimated at the sector level, but we assume that the shock at the sector level is identical for all sectors.

The shocks $\epsilon_{\tau^j_{def,t}}$ and $\epsilon_{GAP^j_t}$ are orthogonalized by introducing the contemporaneous variable $\tau^j_{def,t}$ in the second equation. The system is therefore directly estimated as a structural VAR. This is consistent with the assumption that shocks originate at the microeconomic level and affect the macroeconomy, with feedback effects from the macroeconomic level to defaults.

Our two-equation system at the sector level (j) is therefore[17]:

$$\left\{\begin{array}{l}\tau^j_{def,t}=a_1 GAP^j_{t-2}+b_1 GAP^j_{t-3}+c_1 M_{t-2}+d_1 Z^j_{t-3}+\epsilon^j_{1t}\\ GAP^j_t=b_2 GAP^j_{t-1}+c_2 GAP^j_{t-2}+a_2\tau^j_{def,t}+\epsilon^j_{2t}\end{array}\right. , \quad (3)$$

where a_1, b_1, d_1 and c_1 are estimated from the duration model at the firm level.

The exogenous variables Z^j_{t-3} (financial ratios) are the sector-based aggregate counterpart of the individual financial ratios used in the duration model and M_{t-2} the macroeconomic variables included beyond the output gap, namely inflation, interest rate and exchange rate. ϵ^j_{1t} and ϵ^j_{2t} are orthogonal shocks.

Such a system can be run recursively as a "VAR-X(1) type" model, using $\tau^j_{def,t}$ and χ^j_t with $\chi^j_t = GAP^j_t - a_2\tau^j_{def,t}$:

$$\begin{pmatrix}\tau^j_{def,t}\\ \tau^j_{def,t-1}\\ \tau^j_{def,t-2}\\ \chi^j_t\\ \chi^j_{t-1}\\ \chi^j_{t-2}\end{pmatrix}=\begin{pmatrix}0 & a_1a_2 & b_1a_2 & 0 & a_1 & b_1\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ a_1b_2 & a_1c_2 & 0 & b_2 & c_2 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\end{pmatrix}\begin{pmatrix}\tau^j_{def,t-1}\\ \tau^j_{def,t-2}\\ \tau^j_{def,t-3}\\ \chi^j_{t-1}\\ \chi^j_{t-2}\\ \chi^j_{t-3}\end{pmatrix}+\begin{pmatrix}c_1M_{t-2}\\ 0\\ 0\\ 0\\ 0\\ 0\end{pmatrix}+\begin{pmatrix}d_1Z^j_{t-3}\\ 0\\ 0\\ 0\\ 0\\ 0\end{pmatrix}+\begin{pmatrix}\epsilon^j_{1t}\\ 0\\ 0\\ \epsilon^j_{2t}\\ 0\\ 0\end{pmatrix} \quad (4)$$

Or in a more compact form :

$$X_t = AX_{t-1} + B_{t-2} + C_{t-3} + E_t. \quad (5)$$

We consider shocks to GAP^j_t that are similar across all sectors, e.g. a recession affecting all sectors simultaneously, *i. e.* $\epsilon^j_{1t} = \epsilon_t$, for all j so that we can omit the superscript j. Running equation (5) recursively yields at date $t+l$:

[17] To simplify the presentation of our system, we introduce the macroeconomic variables only with 2 lags. The extension to 3 lags is straightforward.

$$X_{t+l} = A^{l+1}X_{t-1} + \Sigma_{k=0}^{l} A^{l-k}(B_{t+k-2} + C_{t+k-3} + E_{t+k}). \quad (6)$$

Or, in matrix form, after selection of the appropriate components:

$$\begin{pmatrix} \tau^{j}_{def,t} \\ \chi^{j}_{t} \end{pmatrix} = F(L) \begin{pmatrix} \epsilon_{1t} \\ \epsilon_{2t} \end{pmatrix}. \quad (7)$$

IRFs to a positive one standard deviation shock can be computed using standard method on our system. They indicate that the default ratio responds negatively to a shock to the output gap and the output gap responds negatively to a shock to defaults. We provide the standard error around the IRFs using Monte Carlo simulation (with random draws on the $a's$, $b's$, as well as the c and d coefficients).

In Figure 3, notice that the impacts of the shocks are transitory and return to zero after a few years. Based on the IRFs, we can estimate the magnitude of the "second round" effects by the vertical distance between the IRF and the baseline (the horizontal axis at zero). This can easily be understood in the example of a positive shock on the GAP variable. With the default equation, a positive shock to the output gap at year t would only have a negative effect on the default ratio (hence on default frequency) at $t+2$ and $t+3$ (year 3 and 4 in Figure 3). Within our system, the shock on GAP is not a one-time event, because the shock is itself persistent due to its autoregressive structure, but also because lower defaults imply a higher output gap, which decreases further the default ratio. As a consequence, defaults return to baseline only after year 6. Conversely, a negative shock on the output gap would have a more persistent effect, imply a higher default frequency for a longer period than what would predict the LOGIT model.

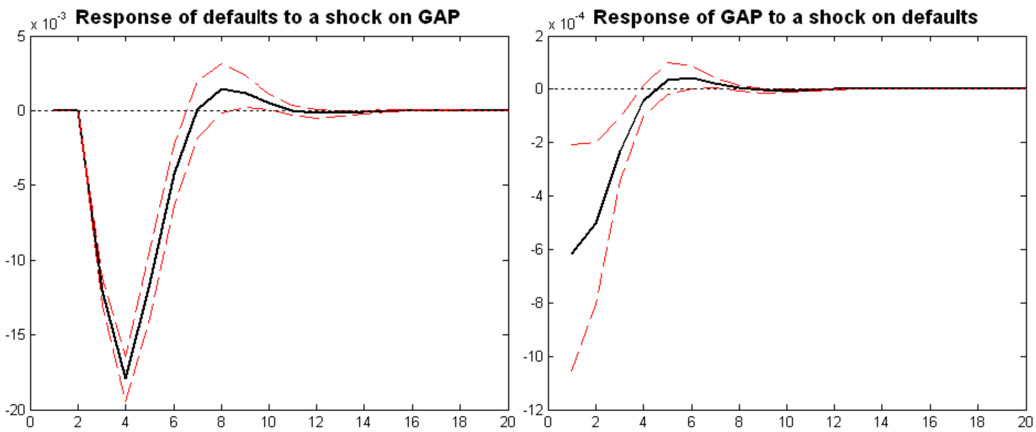

Figure 3 : Impulse Response Function from Panel VAR (%)
(Monte Carlo confidence intervals (90%), 1000 Trials)

4.2 Impulse response analysis integrating the full macroeconomic environment

When running a macroeconomic stress test scenario, one needs also to take into account the effect of a shock on the other exogenous variables than the output gap. This is the component $\Sigma_{k=0}^{l}(A^{l-k}B_{t+k-2}+C_{t+k-3})$ in (6), or simply of the impact of the $Z_t's$ and the $M_t's$. In that respect one should distinguish between the $Z_t's$ and the $M_t's$. The $Z_t's$ may be assumed to be exogenous, consistently with the findings of Carling et *al.* (2003) : macroeconomic variables do not affect individual (and sector-level) financial ratios directly.

Concerning the $M_t's$, however, a shock to ϵ_{2t} has an effect on inflation (INF), the long term interest rate (IRL) and the exchange rate. In order to take this effect into account, we estimate a quarterly VAR.[18] The number of lags is determined by the Akaike criterion. It turns out that the output gap has no significant effect on the exchange rate (see IRF in Figure 2 of Annex C) and we only consider the response of a shock of GAP on inflation and interest rate. We aggregate the response at the annual frequency in order to derive a new path for M_t or $\widehat{B}_{t-2}$. We assume that the following quarterly inverted VAR yields:

$$\begin{pmatrix} INF_t \\ IRL_t \\ GAP_t \end{pmatrix} = H(L)Q^{-1}Q\begin{pmatrix} \eta_{it} \\ \eta_{2t} \\ \eta_{3t} \end{pmatrix} = H(L)Q^{-1}\begin{pmatrix} \eta'_{it} \\ \eta'_{2t} \\ \eta_{3t} \end{pmatrix}. \quad (8)$$

Here, again, Q is an upper triangular matrix (with zeroes on the lower left part) indicating that shocks to interest rate and inflation have no contemporaneous effect on the GAP at the quarterly frequency.

An important issue is also to calibrate the relative size of the shocks affecting the output gap (GAP), interest rates and inflation. We proceed according to the following steps:

- start from the quarterly VAR, and compute the annualized value of the IRFs of inflation and interest rate to a unit shock to the output gap;
- compute the response of a shock of one standard deviation to the innovation of the output gap in the annual/sectorial model, by simply multiplying the previous IRF by one standard deviation of the innovation in the annual/sectorial VAR
- derive the annual path for INF_t and IRL_t which is the new path for B_t that we note $\widehat{B}_t$
- run recursively the equation:

$$\widehat{X}_t = A^t X_0 + \Sigma_{s=0}^{t-1} A^{t-s}(\widehat{B}_{s-2} + C_{s-3} + E_s). \quad (9)$$

[18] The long term interest rate is including in deviation to a quadratic trend in order to take into account the convergence to EMU until 1999. Very similar results are found assuming a linear trend until 1999 and no trend afterwards.

The IRFs for GAP_t^j and $\tau_{def,t}^j$ are shown in Figure 4. They are similar to the ones derived in Figure 3, but they are shifted upwards or downwards due to the effect of the exogenous variables (INF_t and IRL_t through $\widehat{B}_{t+k-2}$). In particular, the impact of a positive shock to the output gap on defaults is slightly less negative than before, due to the positive effect of a shock on the output gap on the long term interest rate, which itself has a positive effect on defaults, as indicated in Table 2. This is consistent with the observation that when the cycle reaches its peak, financially fragile firms are hurt by higher interest rates.

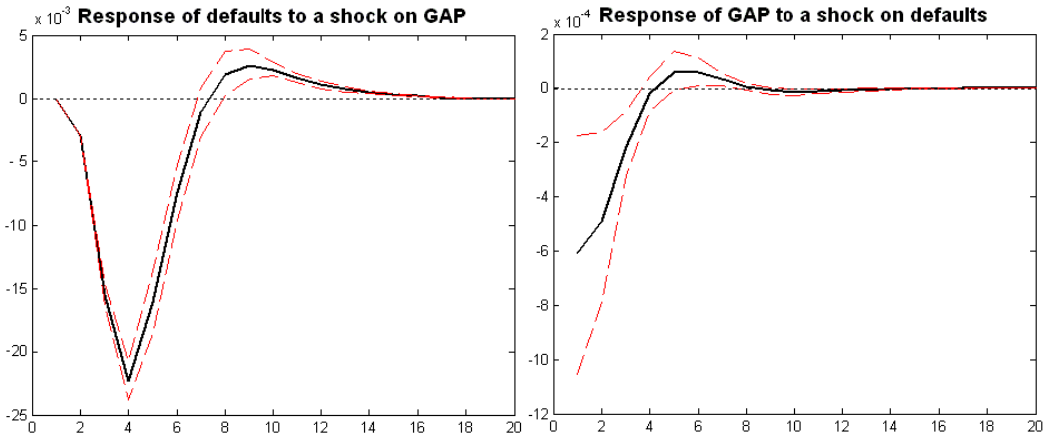

Figure 4 : Impulse Response Function from Panel VAR (%) with additional macroeconomic variables reactions (Monte Carlo confidence intervals (90%), 1000 Trials)

4.3 Impact on the distribution of default probabilities

We describe now how we can assess the effect of stress test scenarios on the distribution of individual default probabilities. We consider here simple scenarios, where we only shock the output gap, taking the other macroeconomic variables as fixed. They correspond to a transitory shock to the default equation in system (3). The first scenario is an increase in the output gap in 2002 (which, instead of decreasing, remains at the level of 2001, where it was equal to zero). The second scenario is a severe recession, measured by a drop in the output gap by 2 standard deviations.

Scenario 1 : More Favorable Business Cycle (GAP in 2002 = GAP in 2001=0)

We consider here a variant scenario where we keep the output gap constant in 2002 at the level of 2001, which implies a more favorable cyclical environment.

This scenario is equivalent to a positive shock on GAP in 2002 with size equal to 0.8122 (= GAP(2001)-GAP(2002)). The following years are unchanged.

The shock shifts the distribution of default probability to the left as a higher fraction of firms experience a lower default probability (see also Figure 5 in Annex D)

Table 4 : Overall Distribution of default probabilities before and after stress 1

(%)	*Mean*	*S.e.*	*Min*	*P*05	*Q*25	*Q*50	*Q*75	*P*95	*Max*
Before stress 1	1.5499	2.3582	3.9×10^{-4}	0.0535	0.352	0.914	1.862	4.946	67.84
After stress 1	1.5039	2.3333	3.9×10^{-4}	0.0529	0.014	0.866	1.860	4.860	67.84

The shift of the distribution of default probabilities year-by-year is more significant than for the overall distribution (*i. e.* for all years, see Table 4). Given that the shock affects default with a certain lag, defaults are mainly affected in 2004 and 2005. We can see from Table 5 (2004) and Table 6 (2005), that the downward shift in mean relative to 2004 (from 1.154% to 1.139%) is more important than the one relative to 2005 (from 0.789% to 0.648%). It corresponds to the larger coefficient of GAP at $t-2$ than $t-3$ in the default equation, as indicated in Table 2.

Table 5 : Distribution of default probabilities in 2004 before and after stress 1

(%)	*Mean*	*S.e.*	*Min*	*P*05	*Q*25	*Q*50	*Q*75	*P*95	*Max*
Before stress 1	1.554	1.844	9.335×10^{-3}	0.214	0.593	1.061	1.858	4.346	54.69
After stress 1	1.139	1.391	6.771×10^{-3}	0.155	0.431	0.772	1.354	3.190	46.68

Table 6 : Distribution of default probabilities in 2005 before and after stress 1

(%)	*Mean*	*S.e.*	*Min*	*P*05	*Q*25	*Q*50	*Q*75	*P*95	*Max*
Before stress 1	0.789	1.048	3.632×10^{-3}	0.103	0.288	0.521	0.926	2.213	2.908
After stress 1	0.648	0.872	2.971×10^{-3}	0.085	0.236	0.426	0.759	1.817	2.512

Scenario 2 : Severe Recession (GAP in 2002 minus two standard deviations)

Here we consider another "stress test" scenario characterized by a severe recession, calibrated by a decrease in the output gap in 2002 by two standard deviations (this corresponds to a negative output gap, which reaches in absolute value the maximum amplitude observed over the 1990-2006 period) instead of the small decline observed in the data. This shock is the reverse of scenario 1, with a larger amplitude. The distribution of default probability shifts therefore to the right. We provide the overall distribution of individual default probabilities in Table 7 (see also Figure 6 in Annex D), which is the equivalent of Table 4. The mean default probability shifts from 1.5% to 1.7%. The median shifts from 0.9% to 1%.

Table 7 : Distribution of default probabilities before and after stress

(%)	*Mean*	*S.e.*	*Min*	*P*05	*Q*25	*Q*50	*Q*75	*P*95	*Max*
Before stress 2	1.550	2.358	3.9×10^{-4}	0.053	0.3524	0.914	1.862	4.946	67.84
After stress 2	1.732	2.556	3.9×10^{-4}	0.054	0.4019	1.029	2.090	5.559	72.65

In order to save space, we do not provide the equivalent of Table 5 and 6, which would also indicate that the change in default, and the associated shift in the default distribution, is more significant in 2004 than in 2005.

5 Conclusion

This paper reports empirical evidence on the links between macroeconomic cycles and changes in financial fragility measured at the microeconomic level by focusing on corporate firms in France.

First, we show that macroeconomic conditions do have influence on the default rate of corporate firms, as proved by the significant impact of lagged macroeconomic variables beside financial ratios in a duration model explaining the life expectancy of non financial firms. Indeed, we find that the output gap included in the duration model with lags of two and three years has a significant negative effect on the default probability.

Second, in the output gap equation, we find a significantly negative coefficient for the observed default rate, which is also a function of lagged financial ratios (and lagged macroeconomic variables for model III).

Third, thanks to a Panel-VAR framework, which allows to take into account the joint dynamic of the output gap and financial fragility, we account for "second round" effects of a shock to the output gap on future defaults, as well as from defaults back to the output gap. This is particularly useful for stress testing.

Further research should investigate different specifications for the default equation as well as various information sets, notably implementing the analysis independently at the sector level. Other specifications for the Panel VAR should also be investigated, in particular a longer lag structure.

References

[1] Allen, L., A. Saunders (2004), "Incorporating systemic influences into risk measurements: a survey of the literature", *Journal of Financial Services Research*, 26 (2), 161-191.

[2] Altman, E. I., (1968), "Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy", *Journal of Finance*, 23, 589-609.

[3] Altman, E. I., (1993), Corporate Financial Distress and Bankruptcy: *A Complete Guide to Predicting and Avoiding Distress and Profiting from Bankruptcy*, John Wiley and Sons, Inc., New York.

[4] Arellano, M. and S. Bond, (1991), "Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations", *Review of Economic Studies*, 58, 277-297.

[5] Arellano, M. and O. Bover, (1995), "Another look at the instrumental variable estimation of error-components models", *Journal of Econometrics*, 68, 385-393.

[6] Bardos M., S. Foulcher and E. Bataille (2004), *Les scores de la Banque de France : Méthodes, résultats et applications*, Observatoire des Entreprises, Banque de France.

[7] Beck N., Katz J. N., Rucker R. (1998), "Taking Time Seriously: Time-Series-Cross-Section Analysis with a Binary Dependent Variable", *American Journal of Political Science*, Vol. 42, N°4, p. 1260-1288.

[8] Beck N. and Rucker R. (1997), "*Conflict in Time and Space*", Center for International affairs Working Paper, N°97-98, Harvard University.

[9] Bernanke, B and M. Gertler (1989), "Agency costs, Net Worth and Business Fluctuations", *American Economic Review*, 79, 14-31.

[10] Blundell, R. and S. Bond (1998), "Initial conditions and moment restrictions in dynamic panel data models", *Journal of Econometrics*, 87, 115-143.

[11] Bonfim D. (2007), "Evaluating the Contribution of Firm Level Information and of Macroeconomic Dynamics", Bank of Portugal *Working papers* 2007-07

[12] Bordes, C and J. Melitz (1992), "Endettement et défaillances d'entreprises", *Annales d'Economie et de Statistique*, 22, oct.-déc.

[13] Campbell J. Y., J. Hilscher, and J. Szilagyi (2008), "In search of distress risk", *The Journal of Finance*, Vol. LXIII, N°. 6, December

[14] Carling, K, T. Jacobson, J. Lindé, Kasper Roszbach (2007), "Corporate credit risk modelling and the macroeconomy", *Journal of Banking and Finance*, 845-868.

[15] Carling, K, T. Jacobson, J. Lindé, Kasper Roszbach (2004), "Exploring relationships between Firms' balance sheet and the macroeconomy" Bank of Sweden, mimeo.

[16] De Graeve, F, T Kick and M. Koetter (2007), "Monetary Policy and Financial (In)Stability : An integrated Micro-Macro Approach", *Journal of Financial Stability* (Forthcoming)

[17] Domens, J (2006), Les défaillances d'entreprises entre 1993 et 2004, DCASPL, *mimeo.*

[18] Drehmann, M. , A. J. Patton and S. Sorensen (2005), Corporate defaults and large Macroeconomic Shocks, mimeo, Bank of England.

[19] Duffie D, Saita L. and Wang K. (2007), "Multi-period corporate default prediction with stochastic covariates", *Journal of Financial Economics*, 83, 635–665.

[20] Eubank, Randall L. (1988), *Spline smoothing and nonparametric regression*, New York: Marcel Dekker

[21] Löffler G. and Maurer A. (2008), "*Incorporating the Dynamics of Leverage into Default Prediction*", Humboldt University, Berlin, Germany in its series SFB 649 Discussion Papers 2009-24.

[22] Lown, C and D. P. Morgan (2006), "The Credit Cycle and the Business Cycle: New Findings Using the Loan Officer Opinion Survey", *Journal of Money, Credit and Banking*, 38:6, 1575-1597.

[23] Misina, M and D. Tessier (2007), The importance of non linearities and macroeconomic uncertainty in stress testing, *mimeo*, Bank of Canada.

[24] Nahmias, L. (2005), "L'impact économique de la défaillance d'entreprises", Bulletin de la Banque de France, N°138, Juin.

[25] Koopman, S. J. and A. Lucas (2005), "Business and default cycles for credit risk", *Journal of Applied Econometrics.*

[26] Shumway, T. (2001), "Forecasting bankruptcy more accurately: a simple hazard model", *Journal of Business*, 74 (1), 101-124.

A Comparison of the likelihood function of the duration model and the multiperiod LOGIT model

Following Shumway (2001), we show here that when the functions F_{LOGIT} and h are related according to the following relationship:

$$h(z'_{(i,t)}\beta) = 1 - F_{LOGIT}(z'_{(i,t)}\beta) = \frac{\exp(z'_{(i,t)}\beta + c)}{1+\exp(z'_{(i,t)}\beta + c)},$$

then the likelihood of the multiperiod LOGIT appears to be the same as the likelihood of the duration model.

First, let us recall the expression of the likelihood for a discrete duration model with a hazard function h :

$$\begin{aligned} L(u_1,...,u_n) &= \prod_{i=1}^{n} h(u_i - \underline{t_i})^{y_i} \prod_{d^{(j)}<u_i-\underline{t_i}} (1-h(d^{(j)})) \\ \text{with } u_i &= Min(t_i, c_i), \end{aligned}$$

and $\underline{t_i}$ (supposed to be positive to avoid left-side censoring) is the date of the start of the activity of firm i; t_i is the date of default; c_i is the first date at which firm i is not observed when it leaves the sample for a cause which is not default (for example, a merger).

Suppose that the hazard function is specified as:

$$h(z'_{(i,t)}\beta) = \frac{\exp(z'_{(i,t)}\beta + c)}{1+\exp(z'_{(i,t)}\beta + c)}.$$

$z_{(i,t)}$ denoting a vector of macroeconomic and microeconomic variables observed at date t or date $t-H$ depending on the lag H introduced to account for delays in the availability of the related information.

The previous likelihood can be rewritten as:

$$L(u_1,...,u_n) = \prod_{i=1}^{n}\left[\frac{\exp(z'_{(i,t_i)}\beta + c)}{1+\exp(z'_{(i,t_i)}\beta + c)}\right]^{y_i} \prod_{t<u_i} \frac{1}{1+\exp(z'_{(i,t)}\beta + c)}.$$

Now, let us consider a multiperiod LOGIT model where the individuals are firm-years and the dependent variable is such that:

$$\begin{aligned} y_{(i,t)} &= 1 \text{ if firm } i \text{ defaults at date } t \\ &= 0 \text{ otherwise,} \end{aligned}$$

with the standard conventions:

$$\begin{aligned} P(y_{(i,t)} &= 1) = F_{LOGIT}(z'_{(i,t)}\beta), \\ \text{with } y_{(i,t)} &= 1 \Leftrightarrow y^*_{(i,t)} < 0 \\ \text{where } y^*_{(i,t)} &= c + z'_{(i,t)}\beta + \varepsilon_{(i,t)}, \end{aligned}$$

and the error $\varepsilon_{(i,t)}$ are distributed as a logistic variable, whose cumulated distribution function is given by:

$$F_{\varepsilon_{(i,t)}}(u) = \frac{1}{1+\exp(u)}.$$

Thus, the likelihood can be written as:

$$L(\left(z_{it}, y_{(i,t)}\right);\ i=1,...,n;\ t=1,...,T) = \prod_{i=1}^{n}\prod_{t=1}^{T}\left[\frac{\exp(z'_{(i,t)}\beta+c)}{1+\exp(z'_{(i,t)}\beta+c)}\right]^{y_{(i,t)}}\left[\frac{1}{1+\exp(z'_{(i,t)}\beta+c)}\right]^{1-y_{(i,t)}}.$$

To account for the fact that certain individuals (i,t) may contribute for 0 to the likelihood, when firm i is not observed at date t, we rewrite the latter likelihood as:

$$L(\left(z_{it}, y_{(i,t)}\right);\ i=1,...,n;\ t=1,...,T) = \prod_{i=1}^{n}\prod_{t<u_i}\left[\frac{\exp(z'_{(i,t)}\beta+c)}{1+\exp(z'_{(i,t)}\beta+c)}\right]^{y_{(i,t)}}\left[\frac{1}{1+\exp(z'_{(i,t)}\beta+c)}\right]^{1-y_{(i,t)}}.$$

The two likelihood functions are therefore identical.

B Estimation Method for the second equation using Panel-GMM

Different GMM estimation methods are available in the literature on panel data which allow for dynamics in the model (*i. e.* presence of lagged dependent variable among regressors), Arellano and Bond (1991), Arellano and Bover (1995) and Blundell and Bond (1998).

To understand the differences across these methods, let us write the dynamic panel regression under the general form:

$$X_{j,t} = \alpha X_{j,t-1} + \beta' \widetilde{Z}_{j,t} + u_j + e_{jt}, \tag{B1}$$

where $\widetilde{Z}_{j,t}$ are other regressors. The residual is decomposed into a fixed individual effect component u_j and a random term e_{jt}.

Arellano and Bond (1991) propose a "first-differenced GMM" estimate which is obtained after first differencing of (B1), to eliminate the fixed effects:

$$X_{j,t} - X_{j,t-1} = \alpha\left(X_{j,t-1} - X_{j,t-2}\right) + \beta'\left(\widetilde{Z}_{j,t} - \widetilde{Z}_{j,t-1}\right) + \left(e_{jt} - e_{jt-1}\right). \tag{B2}$$

The estimation methods usually implemented for static panel models are not adapted in the dynamic case, because the error term $(e_{jt} - e_{jt-1})$ is correlated with $(X_{j,t-1} - X_{j,t-2})$. In the method of Arellano and Bond (1991), the instruments are the lagged variables taken in level, with the moment conditions:

$$\begin{aligned} E\left[X_{j,t-s}.\left(e_{jt} - e_{jt-1}\right)\right] &= 0 \qquad \text{(B3)} \\ E\left[\widetilde{Z}_{j,t-s}.\left(e_{jt} - e_{jt-1}\right)\right] &= 0 \end{aligned}$$

for $s \geq 2$ and $t = 3, ..., T$.

However this estimator has been proved to have a low asymptotic precision. Arellano and Bover (1995) improve the previous estimation procedure by adding moment conditions related to the equation in levels.

Finally, Blundell and Bond (1998) extend the GMM method under relevant hypotheses by combining moment conditions drawn from the equation in levels and from the equation in first differences. Their additional moment conditions can be written as:

$$\begin{aligned} E\left[(X_{j,t-s} - X_{j,t-s-1}) \, . \, (u_j + e_{jt})\right] &= 0 \\ E\left[\left(\widetilde{Z}_{j,t-s} - \widetilde{Z}_{j,t-s-1}\right) . \, (u_j + e_{jt})\right] &= 0 \end{aligned} \qquad \text{(B4)}$$

By Monte Carlo simulation exercises, Blundell and Bond (1998) show that the first-differenced GMM estimate might be biased in small samples and that it is less efficient than the System-GMM estimate in small samples.

Let us note that the estimates can be obtained with either a one- or a two-step procedure. The latter should be better for small samples and this is the one we have implemented. Let us add that two tests are required to validate the estimation method: the overidentification test of Sargan/Hansen which checks that the instruments are not correlated with the error terms and the test for no correlations of order two for the residuals, proposed by Arellano and Bond (1991)[19].

[19] First differencing induces an auto correlation of order one:

$$cov(e_{jt} - e_{jt-1}, e_{jt-1} - e_{jt-2}) \neq 0.$$

C Impulse responses from quarterly VAR

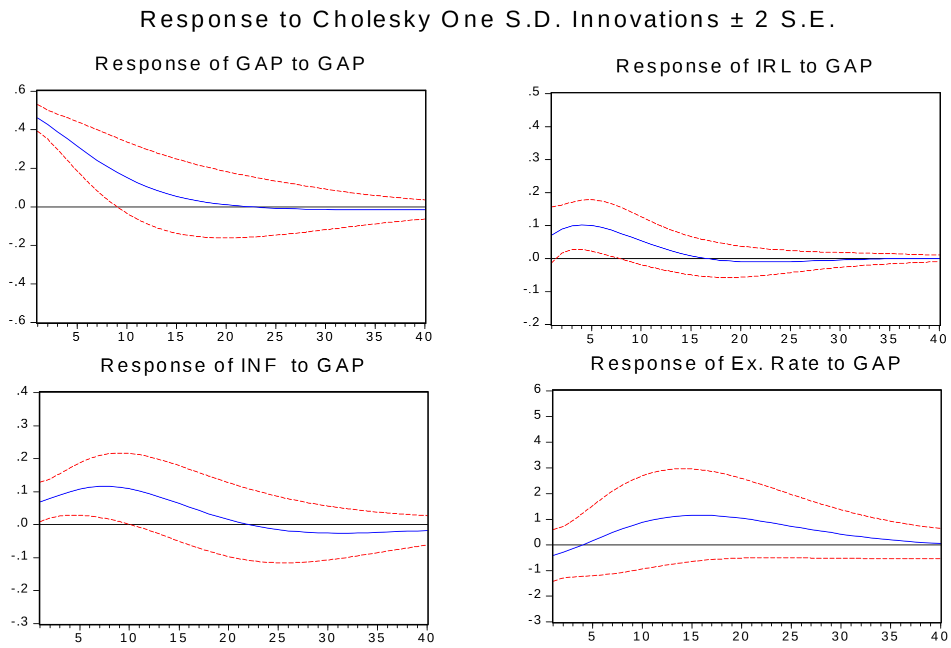

Figure 2 : Impulse responses from quarterly VAR

D Distribution of individual default probabilities

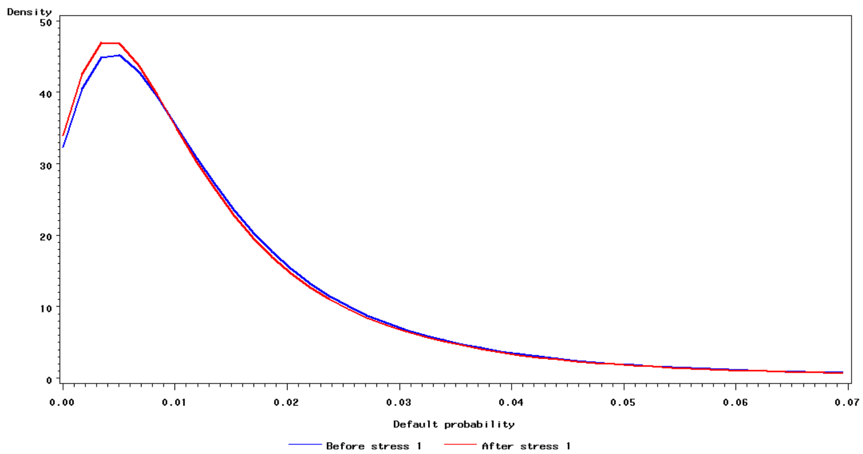

Figure 5: Individual default probability distribution, before and after stress 1

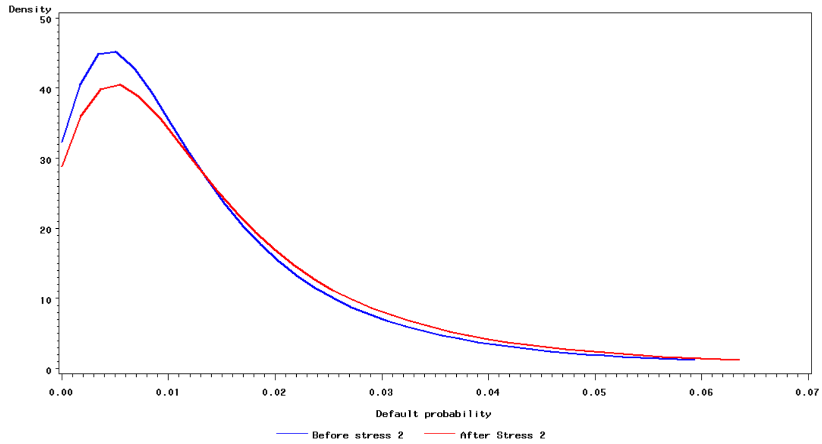

Figure 6: Individual default probability distribution, before and after stress 2

E Estimation results : default at t with information at t-2

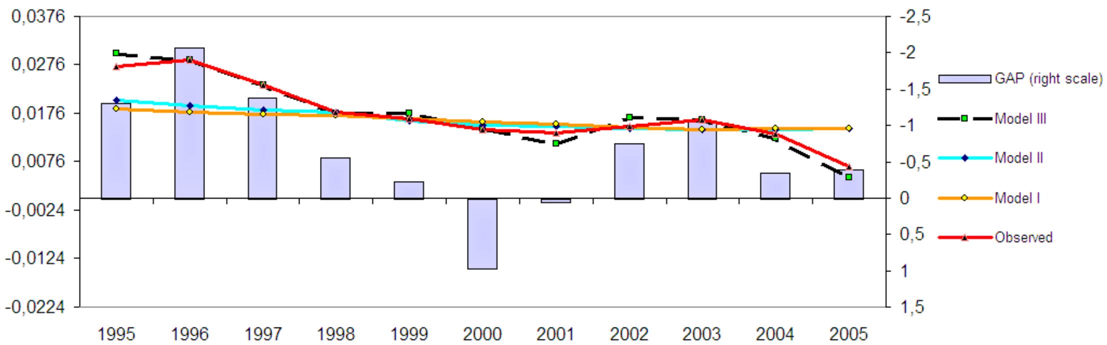

Figure 7 : Aggregate default rate- observed and estimated at t from information at $t-2$ (financial and Macro variables) and $t-1$ (macro economic variables)

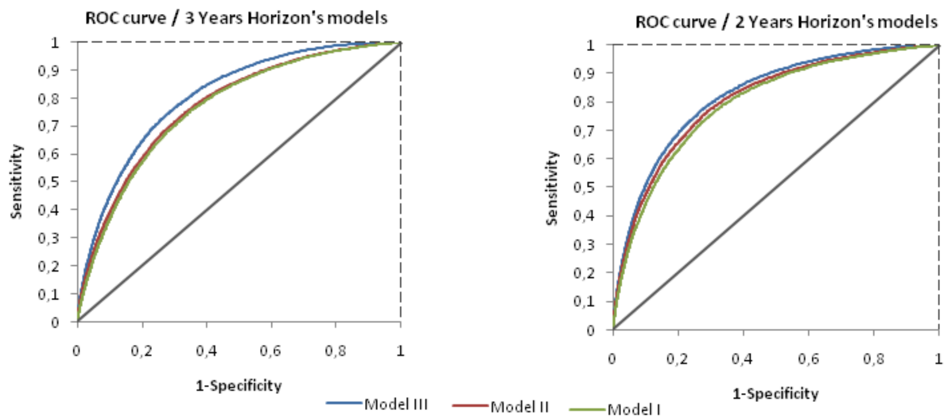

Figure 8 : ROC curves

Table 4 : Logit models estimation of corporate defaults at t / Dynamic approach with information at *t-2*

Variable	*Model I* (*Dyn*) coef.	s.e.	*Model II* (*Dyn*) coef.	s.e.	*Model III* (*Dyn*) coef	s.e.
Micro variables						
$Cst.$	-5.2646^{***}	0.0534	-5.6045^{***}	0.0641	-13.0380^{***}	0.2570
$Profit_{t-2}$	-4.7857^{***}	0.0914	-4.4819^{***}	0.0917	-4.5351^{***}	0.0922
$Leverage_{t-2}$	2.1766^{***}	0.0344	2.0079^{***}	0.0347	2.0041^{***}	0.0350
Liq_{t-2}	-0.0215^{***}	0.0016	-0.0186^{***}	0.0015	-0.0167^{***}	0.0014
$Bus.\ Cred._{t-2}$	0.0018^{***}	0.00018	0.0017^{***}	0.00019	0.00153	0.0002
Int_{t-2}	0.2806^{***}	0.0173	0.2490^{***}	0.0172	0.2082^{***}	0.0171
$Size\ (large)_{t-2}$	-1.7848^{***}	0.0975	-1.6306^{***}	0.0976	-1.4946^{***}	0.0979
$Size\ (medium)_{t-2}$	-0.3396^{***}	0.0397	-0.2784^{***}	0.0399	-0.1518^{*}	0.0402
IP_{t-2}			1.3990^{***}	0.0260	1.3279^{***}	0.0265
$Tax\ Arrears_{t-2}$			0.2919^{***}	0.0407	0.2857^{***}	0.0408
Macro variables						
GAP_{t-2}					-0.1575^{***}	0.0194
INF_{t-2}					-0.3501^{***}	0.0276
IRL_{t-2}					0.3683^{***}	0.0287
$Ex.Rate_{t-2}$					0.0488^{***}	0.0023
GAP_{t-1}					-0.3302^{***}	0.0222
INF_{t-1}					-0.3598^{***}	0.0243
IRL_{t-1}					0.0557	0.0206
$Ex.Rate_{t-1}$					0.0120^{***}	0.0018
$\widetilde{R}^2$ [20]	0.1072		0.1245		0.1446	
LR test$^{(1)}$			443.67^{***}		515.15^{***}	
AUC$^{(2)}$	0.794		0.807		0.823	
Nb. obs.	876,999		876,999		876,999	

Notes : ***,** and * denote significant at 10, 5 and 1% levels, respectively.; $^{(1)}$Likelihood Ratio test statistic for nested models $\sim \chi^2_{(df)}$; $^{(2)}$Area Under the Receiver Operating Characteristics (ROC) Curve; with $df = 2$ (*resp.* 8) *for Model I vs II* (*resp. Model II vs III*)

[20] Nagelkerke's $\widetilde{R}^2$ is obtained as: $\frac{R^2}{R^2_{\max}}$ with $R^2 = 1 - \left\{LnL_{\max}(\beta=0)/\,LnL_{\max}(\widehat{\beta})\right\}^{\frac{2}{N}}$ and $R^2_{\max} = 1 - \{LnL_{\max}(\beta=0)\}^{\frac{2}{N}}$. $\widetilde{R}^2$ measure the discrepancy between the LOGIT models with $(\widehat{\beta})$ or without $(\beta = 0)$ the dependent variables X. N is a sample size.

F In and Out of sample Results

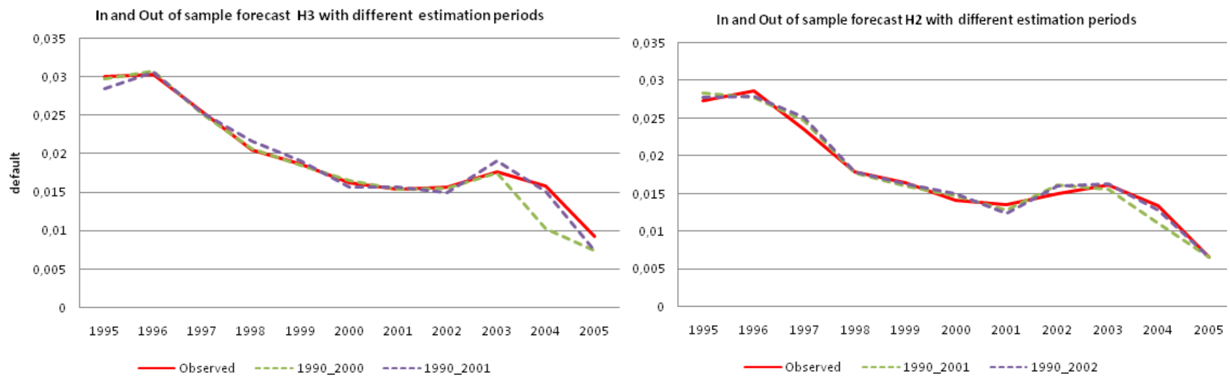

Figure 9 : Performance In and Out of sample of Model III (H3 and H2) **without** spline method

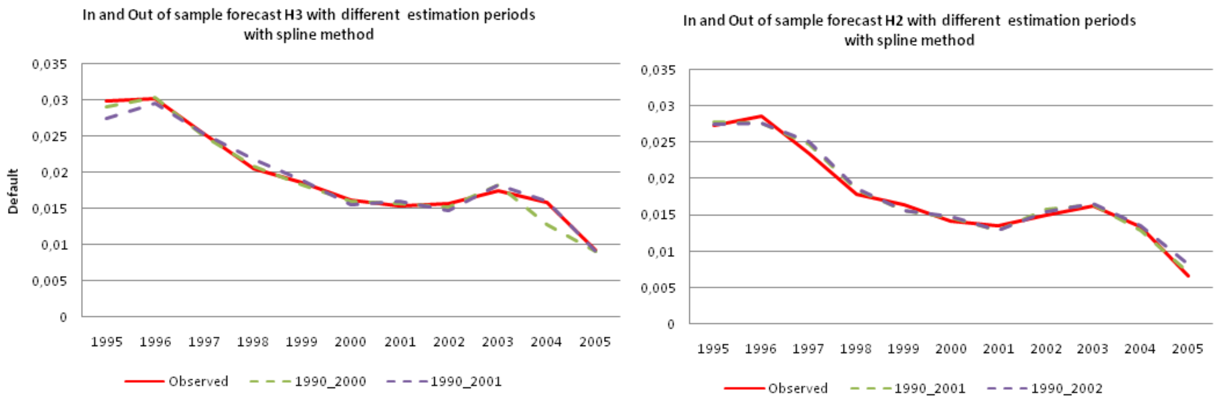

Figure 10 : Performance In and Out of sample of Model III (H=3 and H=2) **with** spline method

G Summary statistics

Table A: Descriptive statistics for three years horizon data by firm type

Variable	Mean	S.d	Min	25%	50%	75%	Max
Non-defaulted firms ($N = 849,628$)							
Profitability	0.1187	0.0996	-0.1975	0.0601	0.1048	0.164	0.532
Leverage	0.5528	0.2088	0.00000	0.4017	0.5498	0.699	1.195
Liquidity	4.9986	21.030	0.00000	0.0440	0.3211	1.563	172.9
Bus.Cred.	70.654	41.548	0.00000	42.763	65.634	89.33	300.0
Interest	0.1178	0.2972	-1.8372	0.0233	0.0769	0.175	2.250
Defaulted firms ($N = 13,377$)							
Profitability	0.0716	0.1098	-0.1975	0.0204	0.0689	0.122	0.532
Leverage	0.7195	0.2015	0.06664	0.5906	0.7374	0.855	1.195
Liquidity	1.8660	11.879	0.00000	0.0093	0.0738	0.376	172.9
Bus.Cred.	84.059	46.481	0.00000	54.038	77.304	103.2	300.0
Interest	0.1796	0.5115	-1.8372	0.0460	0.1724	0.313	2.250