www.cefin.unimore.it

# CEFIN Working Papers
No 25

## Modelling credit risk for innovative firms: the role of innovation measures

by Chiara Pederzoli, Grid Thoma, Costanza Torricelli

March 2011

CEFIN – Centro Studi di Banca e Finanza
Dipartimento di Economia Aziendale – Università di Modena e Reggio Emilia
Viale Jacopo Berengario 51, 41100 MODENA (Italy)
tel. 39-059.2056711 (Centralino) fax 059 205 6927

**Modelling credit risk for innovative firms: the role of innovation measures**

Chiara Pederzoli[♠♣], GridThoma[♦♣], Costanza Torricelli[★♣]

[♠] University of Milano-Bicocca
[♦] University of Camerino
[★] University of Modena and Reggio Emilia
[♣] CEFIN, Centro Studi Banca e Finanza, Modena

## Abstract

Financial constraints encountered by small-medium enterprises (SME) are particularly severe for innovative firms, which, in the EU, cannot rely on a sufficiently developed venture capital industry and have to depend on debt capital. It is thus important to develop models which, in consideration of the specific features of innovative SMEs, provide a reliable estimate of their probability of default (PD) that can also serve as a rating of the innovative firm. Based on the signaling value of innovation-related assets such as patents, this paper shows the role of innovative assets in credit risk modeling. Specifically, we include in a logit model two innovation-related variables in order to account for both the dimension and the value of the patent portfolio. Based on a unique dataset of innovative SMEs with default years 2006-2008 we show that, while the value of the patent portfolio always reduces the PD, its dimension increases the firm's riskiness unless coupled with an appropriate equity level.

*Keywords:* innovative SMEs, default probability, patent value, R&D productivity

*JEL classification:* G21, G32, C25, O31, O34

*Corresponding author*: Costanza Torricelli, Dipartimento Economia Politica, University of Modena and Reggio Emilia, Via Berengario, 51 - 41121 Modena – Italy, Tel. +390592056733, Fax +390592056937, e-mail: costanza.torricelli@unimore.it

## 1 Introduction

The advent and the fast growth of the knowledge economy and the parallel development of science-based industries (e.g. biotechnology, software) have been accompanied by the emergence and success of innovative start-ups, which in many instances have outperformed incumbent firms. Examples include Microsoft in operating systems, Google and Yahoo in web applications, Amgen and Genentech in biotechnology, Echelon in automation and many others. This evidence could be directly related with the higher experimentation and innovation propensity of small firms vis-à-vis large ones (Arrow (1975), Arora et al. (2001), Harhoff, 1996).

Innovative firms, independently of their size, face financial constraints as stressed by a broad literature, which has given special attention to the role of equity finance (Carpenter and Petersen, 2002; Brown et al., 2009). Hall (2002) concludes that "… the capital structure of R&D-intensive firms customarily exhibits considerably less leverage than that of other firms". The issue has been extensively surveyed in Hall and Lerner (2010), who claimed that financial constraints are fuelled by information asymmetries between inventor/entrepreneur and investor. In particular, these asymmetries regard the fact that an inventor has a better understanding on the potential success and structure of the R&D project, and thus, the marketplace for financing the innovative assets could be characterized by a typical "lemons market problem". They also stress that financial constraints are particularly severe for R&D projects developed by SMEs, which - as it is the case for non innovative firms too – do not normally rely on equity markets. At the same time, innovative SMEs encounter a stronger adverse selection in the credit market (Harhoff and Körting, 1998): since new innovators are corresponded by a financial distress in an early stage, they face comparatively higher interest rates and reduced credit availability, which in turn have an effect on their financial performance. This hampering mechanism is even more pronounced in sectors and/or countries where the venture capital (VC) industry is not sufficiently developed, as in the EU countries (EIB, 2009). With the exception of UK, the continental EU countries show very low intensity of VC investments relatively to their GDP compared to USA, Israel, Canada and Switzerland (OECD, 2009b).

In sum, innovative SMEs add to the well-known financing difficulties of "traditional" SMEs the above-mentioned problems typical of innovative firms, thus encountering peculiar difficulties in financing their activities. If innovative SMEs have to rely also on debt capital, it is particularly important to develop models which, in consideration of their specific features, provide estimates of their probability of default (PD) and can provide at the same time a rating of the firm. The issue is relevant also in terms of capital regulation given that the Basel Committee on Banking Supervision

(BCBS) recognizes a different treatment for the exposures towards SMEs, which, since the advent of Basel II, benefit from a reduction in the capital requirement proportional to their size.

As for the research on this topic, on one hand there is a broad empirical literature on SMEs default prediction and, on the other, there are many research works in entrepreneurial finance. The former proves, over different period and different datasets, the good performance of logit/probit models[1] and, despite some differences among various research works, a convergence emerges on five categories of financial indicators (leverage, liquidity, profitability, coverage, and activity), and the importance of qualitative variables is also recognized (e.g. Grunert et al., 2005). The latter highlight that patents can constitute a rich information source for financial investors in assessing the quality of innovative firms.[2] Hsu and Ziedonis (2008) show that patents improve the terms by which new firms access venture capital. In particular, they document that the larger the patent portfolio of start-ups, the bigger the money evaluation by VCs and that this effect is even more pronounced for younger and inexperienced firms. In the same vein, Harhoff et al. (2009) demonstrate similar findings and argued that the granting decision by the patent office does not trigger additional financial evaluation from VCs because this event is fully anticipated thanks to information indicators revealed in the patent application (e.g. such as patent citations). A few recent papers, e.g. Buddelmeyer et al. (2010) and Motohashi (2011), estimate hazard or binary choice models to investigate the relation between innovation and company survival. However, none of these works aims to develop and test a proper model for the estimation of a PD.

This paper rests on the most used binary regression models for the estimation of the SMEs default probability and has a twofold aim. First, we test whether the credit quality of innovative firms is better predicted when, beside indicators related to the balance-sheet, the model includes variables reflecting the patent portfolio. Second, we aim to disentangle the different roles played by the dimension of the patent portfolio, its value and the capital structure of the firm.

To this end we use a logit model to estimate a PD on a unique and novel dataset of innovative SMEs with default years 2006-2008. We begin with a standard specification of the model relying on balance-sheet variables only. Then, to test whether the dimension of the patent portfolio reduces the riskiness of innovative firms, we add a variable reflecting the R&D productivity. In the light of the results obtained, we introduce a second innovation-related variable to assess the role played by the value – beside the dimension - of the patent portfolio. Finally, we test an explanation of our findings

[1] For a broader discussion of the issue, see Altman and Sabato (2007).

[2] For a survey on developments of entrepreneurial finance see Denis (2004).

that rests on the capital structure of the firm. As far as we know this is the first attempt to consider jointly financial ratios and innovation measures to predict the PD of a firm.

The paper is organized as follows. Section 2 describes the original dataset. While Section 3 recalls the motivation for the default prediction model used, Section 4 illustrates the specific issues connected with the measurement of the innovation-related regressors. Section 5 discusses the results obtained and the last Section concludes. The Appendixes report descriptive statistics of the sample and of the variables entering the regressions.

## 2 Data description: the sample of innovative firms

In this paper we use a unique and novel dataset of innovative SMEs based on PATSTAT database, EPO BULLETIN and AMADEUS (Thoma et al. (2010) for more details). The first issue to be addressed in the construction of the dataset is the definition of innovative firms. To this end, we use patent data: while not all inventions are patented, patenting activities have increased significantly in the last decade in terms of larger company patent portfolios and larger share of firms applying for patents in many different technologies (OECD, 2009a). On the other hand, patents can be considered a highly objective data source over time and they provide very detailed information regarding the invention and its inventors (Griliches, 1990).

To define the set of innovative companies in this paper we include all European firms that have filed at least one patent application in the EPO and PCT/WIPO system.[3] We decided to limit our analysis to these two patent systems in order to take into account only the most relevant patent inventions by a firm and to achieve higher homogeneity across patent measures. In fact, patents document varies significantly in terms of their economic value depending on the legislation (OECD, 2009a).

The data source is the PATSTAT database (version April 2009) and EPO BULLETIN (version December 2009).[4] In particular, our database covers all patent document publications –applications and grants – since the inception of EPO and PCT/WIPO system up to Dec 31st, 2009. Then, relying on the AMADEUS business directory we integrate the patent owner names with demographic and

[3] EPO is the acronym for European Patent Office, whereas PCT/WIPO for Patent Cooperation Treaty/World International Patent Office. For information on these patent systems see Guellec, and van Pottelsberghe de la Potterie (2007).

[4] Both the PATSTAT and BULLETIN database are available to any user under request from the EPO. The data have been managed by with SQL and STATA software toolboxes. For more details on this task see Thoma et al.(2010).

accounting information, such as sector activity, ownership, balance sheet, profit and loss account. Our methodology builds over a previous contribution by Thoma et al. (2010), who developed a complex matching algorithm to merge extensive company information. Our dataset relies on the overall population of patent owners, which allows to overcome any selection bias limitation.

Given the focus on SMEs, the second issue concerns the definition adopted to identify this category of firms. The definition given by the European Union refers both to the number of employees and to sales: firms are considered small if they have less than euro 50 million in sales or less than 250 employees.[5] The Basel regulation for the purpose of capital requirements imposes a criterion based on sales only to discriminate between SMEs and corporates: firms with annual sales less than 50 million euros are considered SMEs.[6] In our sample, we have included firms with turnover in the range of 1-50 million Euros, whereas the geographical context has regarded EU15 countries, Switzerland and Norway.[7] Thus, consistency with the Basel regulation allows to use the estimated PDs as input in capital requirement formula.

A further important issue is the definition of default to be used to classify defaulted firms in our sample and the literature does not provide a univocal one. Altman and Hotchkiss (2006) stress that four terms - *failure, insolvency, default* and *bankruptcy* - are used interchangeably in the literature but have different meaning and refer to different situations in different countries' bankruptcy law. The Basel regulation (BCBS, 2006) adopts a wide default definition in that "a default" is considered to have occurred with regard to a particular obligor when either or both of the two following events have taken place:

- The bank considers that the obligor is unlikely to pay its credit obligations to the banking group in full, without recourse by the bank to actions such as realizing security (if held).
- The obligor is past due more than 90 days on any material credit obligation to the banking group overdrafts will be considered as being past due once the customer has breached an advised limit or been advised of a limit smaller than current outstandings.

---

[5] Commission Recommendation 96/280/EC of April 3, 1996, updated in 2003/361/EC of May 6, 2003. See http://europa.eu/scadplus/leg/en/lvb/n26026.htm.

[6] This implis for the intermediary a reduction in capital requirement proportional to the firm's size The reduction applies to the capital function through the correlation, which is reduced by a maximum of 0.04 for the smallest firms. This correction is justified by the assumption that defaults of small firms are less correlated and therefore less risky on the whole for the portfolio.

[7] The turnover is given by the sum of sales and net stocks of the reference year. In the present analysis we use turnover and not sales because the AMADEUS does not report information on sales for some countries such as UK, Ireland and Denmark.

Often default definitions for credit risk models concern single loan defaults of a company versus a bank, as also emerges from the Basel instructions above. This is the case for banks building models based on their portfolio data, that is relying on single loans data which are not public (e.g. Altman and Sabato (2005)). However, traditional structural models (i.e. Merton-type models) refer to a firm-based definition of default: a firm defaults when the value of the assets is lower than the value of the liabilities, that is when equity is negative.

In this work, we identify a firm's legal status according to the following taxonomy:

i) Active: if a company is currently performing economic activities;
ii) Inactive: if a company has not been performing economic activities in the last three years;
iii) Bankrupted: (a) unable to pay the creditors; (b) the assets are held by a receiver; (c) assets and property of the company redistributed;
iv) Dissolved: when the legal life of company has come to an end;
v) Merged-demerged-acquired: whether a firm has been merged with another company, acquired or split in more than one other company;
vi) Unknown, firms with unavailable legal status.

Consistently with previous studies (e.g. Altman and Sabato, 2007) we include only firms with a legal status active or bankrupt. The reason for this choice lies in the data availability but it is also motivated by the objective of the paper: our aim is to define a model, based on public and accessible data, that measures the health state of the firms and enables any economic subject interested in a specific firm's health (i.e. suppliers, customers, lenders, etc.) to estimate the probability of a particular firm to get bankrupted.

In line with previous literature we adopted a reduced sample approach with a ratio of bankrupt firms of 6% of the overall sample. This rate is the sample default rate before cleaning and is in line with the one assumed by Altman and Sabato (2007). To build our sample, we start with all bankrupt firms with available information on profit/loss and balance sheet accounts in five macro business activities: low tech process industries (US SIC 10-33), chemicals and pharmaceuticals (US SIC 28), manufacturing (US SIC 34-39), distribution (US SIC 50-60) and services (US SIC 70-99). Then, we randomly select firms with active legal status up to 94% of the sample and, in order to obtain a full independence of the observations, we adopt a sampling strategy without replacement.

Finally, we adopt a pooled cross section logit model – as described in the next Section – to estimate a PD with default years 2006-2008 that correspond respectively to fiscal years 2004-2006. The final sample consists of 2,665 firms, whereby 160 are classified as default and 2,505 are active companies according to the AMADEUS business directory.

The sample descriptive statistics are in the Appendix A. Table A1 reports the distribution of firms by cohort and macro industry. To be noticed that about 2/3 of our firms originate from manufacturing and low tech industries, whereas Services account for only 13.7%; the age distribution of the firms is relatively old, with about 70% of the firms incorporated before year 1990. Similarly, Table A2 represents the 160 defaults firms by cohort and macro industry: we can notice the that this distribution follows evenly the statistical patterns of Table A1.

## 3. The default prediction model for innovative SMEs

There is a wide range of default prediction models, i.e. models that assign a probability of failure or a credit score to firms over a given time horizon. The literature on this topic developed especially in connection with the discussion on Basel II, which allows banks to set up an internal rating system, that is a system to assign ratings to the obligors and to quantify the associate PDs. However, some sophisticated models available in the literature can be used only if market data on stocks (structural models) or corporate bonds and asset swaps (reduced-form models) are available. As for SMEs, for which market data are generally not available, either heuristic (e.g. neural network) or statistical models can be applied. Beaver (1966) and Altman (1968) first used discriminant analysis (DA) to predict default. In order to overcome the limits inherent in DA (e.g. strong hypotheses on explanatory variables, equal variance-covariance matrix for failed and not failed firms), since the seminal paper by Ohlson (1980) logit and probit models have been widely adopted.[8] An important advantage of the latter models is the immediate interpretation of the output as a default probability. Focusing on SMEs, a few recent works use logit/probit models, or some evolution of the same, for the PD estimation: Altman and Sabato (2007) use a dataset of US SMEs, Altman and Sabato (2005) analyse separately US, Australian and Italian SMEs, Behr and Güttler (2007) and Fantazzini and Figini (2009) analyse German data, Fidrmuc and Heinz (2009) use data from Slovakia, and Pederzoli and Torricelli (2010) focus on the Italian case. Despite some differences among these analyses, a convergence emerges on a few types of financial indicators, which can be grouped into five categories: leverage, liquidity, profitability, coverage, and activity.

[8] A number of papers, among which Lennox (1999) and Altman and Sabato (2007), show that probit/logit models outperform DA model in default prediction.

Thus, in line with most of the literature on SMEs, we use a binary logistic regression model to estimate the default probability: we quantify the dependent variable according to the definition of default given in Section 2, while we consider both balance-sheet variables and innovation-related variables as regressors.

In the case under investigation in this paper, i.e. innovative SMEs, one issue is still the selection of appropriate and informative balance sheet variables, but the main one is the definition and the measurement of the innovation-related regressors. While the former is tackled by means of a standard backward elimination procedure based on the Schwartz Information Criterion (SIC) and is recalled in Section 5, the latter requires discussing specific issues as illustrated in the following Section.

## 4. The innovation-related regressors

An increasing number of studies use patent counts and other patent-related indicators to measure the quantity and the value of inventive outputs. Several studies show that patent counts are strongly correlated to size of innovative investments typically measured by R&D (e.g. Griliches, 1990). However, crude patent counts are a biased indicator of inventive output because they do not account for differences in the value of patented inventions. This is the reason why innovation scholars introduced four main patent-related indicators as a measure of the value of the inventive output. First, the number of inventors of a patent is associated to the economic and technological value of patents: the technical value of an invention is related to the research cost of the underlining R&D project, which is made up in large part of wage bills for the human resources involved in the project (Harhoff and Thoma, 2010). In this direction, the more inventors in a patent, the more research-intensive and expensive the R&D project (Guellec and van Pottelsberghe (2000), Gambardella et al., 2008). A second indicator is given by the geographical scope of patent protection, i.e. the number of national and international offices in which a patent document has been applied. Typically, the international patent protection requires additional filling costs and this decision by the owner signals a higher expectation of economic value related to the invention (Lanjouw, and Schankerman, 2004; Hall et al, 2007). Third, the number of citations received (henceforth also forward citations) is widely used as indicator of patent value (Harhoff et al.(2003), Trajtenberg, 1990). The literature provides two main explanations: on one hand it demonstrates the cumulativeness of a given technology, suggesting additional R&D being performed and hence market potential, on the other, since citations reveal a knowledge transfer process, it shows that a technology is being used and hence it is valuable. Fourth, the number of technological classes has been shown to be an indicator of technological value similar to the number of citations by Lerner (1994). In particular the number of International Patent Classifications (IPC) codes can be viewed as a measure of technological scope or generality of the patent.

In the empirical analysis of the present paper, we follow Hall et al (2001), who proxy the knowledge assets of a firm following two directions, and we define two innovation-related variables: the former aims to represent the normalized dimension of the patent portfolio, the latter its value. Specifically, the first is defined as the capitalized patent counts standardized by the R&D stocks and can be interpreted as the R&D productivity at the level of the firm. Due to the lack of information on the R&D expenditures for SMEs we measure the patent productivity as the ratio of patents counts divided by the number of active inventors over a 5 years window. Previous research (Harhoff and Thoma, 2010) shows that R&D investments are made up of about 70% of labor costs – typically wage bills for the R&D personnel– and the remaining part is highly correlated with the size of the R&D personnel. Moreover, we think that for the SME case this measure is more suitable than the one based on R&D investments because even when R&D investment is reported in the P&L account it may underestimate the actual intensity of innovation activities. Indeed, in the case of SMEs, R&D activities are not formalized in structured labs and typically R&D costs are mixed with labor costs and/or with other fixed costs when R&D is outsourced. To define the second variable, which captures the value of the inventive output, we built a multidimensional factor index according to the methodology explained in the next Section.

### 4.1 Patent value factor index

Three indicators – family size, citation and IPC technical classes[9] – are combined into a composite index of patent value derived from a common factor model, following the approach suggested by Lanjouw and Schankerman (2004). We use a multiple-indicator model with an unobserved common factor:

$$y_{k\,i} = \lambda_k q_i + \beta' \boldsymbol{X} + e_{ki}$$

where $y_{ki}$ indicates the value of the $k$th patent indicator for the $i$th patent; $q$ is the common factor with factor loadings $\lambda_k$ and normal distribution, and $\boldsymbol{X}$ is a set of controls. The main underlying assumption is that the variability of each patent indicator in the sample may be generated by the

[9] To guarantee a reasonable level of precision, we use the number of eight-digits IPC classification codes reported in the patent document.

variability of a common factor across all the indicators and an idiosyncratic part $e_k$ with distribution $N(0,\sigma^2_k)$, not related to the other indicators.

In our setting, the common factor is the unobserved characteristic of a patent that positively influences three value indicators. Estimation of common factor index $q$ is based on information extrapolated from the covariance matrix of our three indicators. By assuming the normality of $q_i$ and $e_k$ we can estimate by maximum likelihood, which ensures a unique solution. Once the estimates of $\lambda_k$ are obtained, the model is inverted to calculate $q$.

### 4.2 The depreciation problem

One key aspect of knowledge is cumulativeness, that is the knowledge assets of a firm strongly depend on previous vintages of other knowledge. However, knowledge depreciates too and the pace of this process is more fierce in some areas than others. For example, in the last years the rate of technical change in software and other ICTs related industries has been considered very fast.

In the literature to account for time dimension of the knowledge accumulation process previous contributions have adopted conventional declining balance formula using a directly comparable relation with ordinary investment and capital:

$$K_t = R_t + (1-\delta)K_{t-1}$$

where $K_t$ is stock of knowledge at time $t$, $R_t$ the production of knowledge between $t-1$ and $t$, and $\delta$ is the depreciation rate. Although a variety of choices for the depreciation rate have been explored in the past, the choice makes little difference for estimation, and most of previous works use the 15 per cent (see for a survey Hall (2005) and Hall, Mairesse and Mohnen, 2010).

For R&D investments or personnel, typically the starting stock is calculated for each firm at the first available R&D observation year as $K_o= R_o/(\delta+g)$, where $g$ is a conventional growth rate, and approximated with 8 per cent. This assumes that real R&D has been growing at a constant annual growth prior to the sample. Similarly, patent-related variables are obtained using the same method. However, given the longer pre-sample history of patenting (back to 1970s) than for R&D the impact of the initial stock is minimal and thus the initial available patent counts are often not discounted to obtain an initial capital stock.

### 4.3 Data censoring and other measurement issues

Patent data suffer several truncation issues. First, EPO/PCT patent application information are available only with a time delay. A patent application is generally published 18 months after it was filed, whereas the time lag between filing and grant or refusal of patents is not fixed. In our analysis to overcome this end of sample bias we considered all the patent applications and not just grants. Second, the filing date cannot always be defined as the closest recorded date to the invention activity if the EPO/PCT patent application is secondary filing of a priority patent from a national office – and typically this is the case. Hence, we considered as reference year for the patent information the priority year rather than the application year. A third censoring problem regards the patent value indicators. In particular, forward citations to a patent take place over a very long period of time. Based on the empirical evidence (Hall et al., 2007) and given that our firm sample regards fiscal year 2004-2006, we opt to count the forward citations only those taking place after three years from the priority date in order to achieve a homogenous measure across years. Another measurement problem of the patent value indicators concerns the different statistical structure across technologies. For example, citations cumulate more slowly in Chemicals rather than Electronics, because the pace of technical change is faster in the latter technologies. In turn family size in globalized industries such as Pharmaceuticals is higher than Mechanicals. Similarly, number of IPC classes is more numerous in general purpose technologies such as ICTs rather than in Consumer goods. In the literature, there are several statistical procedures to correct for this bias (Hall et al., 2001), but the most frequent approach is detrending by time and technology fixed effects. In this work we scaled our three indicators by the geometric averages computed by reference year and technology groups.[10]

## 5. Model estimation and results

In order to prove the importance of including innovation-related variables and to highlight their differential role, we estimate different variants of the model. Table 1 summarizes the variables

[10] We followed the technology grouping proposed by (OST, 2006) which is articulated in 30 categories. In particular: 1 Electrical devices - electrical engineering; 2 Audiovisual technology; 3 Telecommunications; 4 Information technology; 5 Semiconductors; 6 Optics; 7 Analysis, measurement, control; 8 Medical engineering; 9 Nuclear engineering; 10 Organic fine chemicals; 11 Macromolecular chemistry, polymers; 12 Basic chemical processing, petrol; 13 Surfaces, coatings; 14 Materials, metallurgy; 15 Biotechnology; 16 Pharmaceuticals, cosmetics; 17 Agriculture, food; 18 General processes; 19 Handling, printing; 20 Material processing; 21 Agriculture & food machinery; 22 Environment, pollution; 23 Mechanical tools; 24 Engines, pumps, turbines; 25 Thermal techniques; 26 Mechanical elements; 27 Transport; 28 Space technology, weapons; 29 Consumer goods & equipment; 30 Civil engineering, building, mining.

entering the model. Appendix B reports the descriptive statistics of the variables, by industry, for the whole sample.

Table 1. Variables included in the prediction model

| *Variable name* | *Variable description* |
|---|---|
| *Dependent variable* | |
| DEFAULT | Dummy variable which takes value 1 if a firm is bankrupted: (a) unable to pay the creditors; (b) the assets are held by a receiver; (c) assets and property of the company redistributed. |
| *Independent variables* | |
| *Accounting* | |
| EQ_RAT | The equity ratio of the firm equals Equity / Total Debt. |
| LIQ_RAT | The liquidity ratio is given by Cash/Sales. |
| PROF_RAT | The profit ratio is given by Net Earnings / Total Assets. |
| EX_RAT | This ratio is given by Retained Earnings/Total Assets. |
| COV_RAT | The coverage ratio is given by EBITDA/Interest expenses |
| SALES | The sales variable is measured by the log of Operative Turnover of the firm. |
| *Innovation-related variables* | |
| INV_RAT | The R&D productivity ratio equals Capitalized Patent Stock / Capitalized R&D personnel. |
| VAL_RAT | The patent value ratio equals Capitalized Patent Value Stock / Capitalized Patents Stock. We include three measures of patent value: i) forward citations; ii) size of the patent family; iii) number of patent classes. |
| *Control variables* | |
| Quoted dummy | Firm traded in the stock market. |
| Country dummies | Macro areas: Central Europe (AT, CH, and DE); Benelux (BE, LU and NL); Nordic countries (DK, FI, IS, NO, and SE); Central-South Europe (ES, FR, GR, IT and PT); and others (GB and IE). |
| Cohort dummies | pre 1970, 1970s, 1980s, 1991-1995, 1996-2000, and post 2000. |
| Sectorial dummies | process industries, manufacturing, chemicals and pharmaceuticals, utilities, distribution and retail, and services. |
| Year dummies | 2006, 2007, and 2008. |

Table 2 reports the results of the logistic regression analysis for the in-sample dataset. [11] We begin with a standard specification of the model relying on accounting variables only, which are selected among the same candidate predictors proposed in Altman and Sabato (2007) by means of a backward elimination procedure relying on the Schwartz Information Criterion (SIC). The selected accounting variables, which are meant to describe the five main features of a firm's profile as recalled in Section 3, are: leverage (EQ_RAT), liquidity (LIQ_RAT), profitability (PROF_RAT and EX_RAT), coverage (COV_RAT) and activity (SALES). As expected, all the coefficients have negative sign and are significant (Model 1).

Table 2. Multivariate logistic regression: results

| | *Model 1* | *Model 2* | *Model 3* | *Model 4* | *Model 4 Elasticity* |
|---|---|---|---|---|---|
| SALES | -0.298 *** (0.090) | -0.291 *** (0.090) | -0.287 *** (0.090) | -0.287 *** (0.090) | -0.002 |
| EQ_RAT | -0.665 *** (0.265) | -0.658*** (0.265) | -0.656 *** (0.266) | -0.703 *** (0.239) | -0.006 |
| LIQ_RAT | -3.320 *** (1.317) | -3.194 ** (1.319) | -3.241 ** (1.332) | -3.254 *** (1.330) | -0.028 |
| PROF_RAT | -0.651 ** (0.297) | -0.691 ** (0.296) | -0.704 ** (0.294) | -0.774 *** (0.284) | -0.007 |
| EX_RAT | -1.404 ** (0.690) | -1.454 ** (0.696) | -1.493 ** (0.737) | -1.654 ** (0 .714) | -0.014 |
| COV_RAT | -0.027 *** (0.006) | -0.026*** (0.006) | -0.027 *** (0.006) | -0.026 *** (0.006) | -0.022 |
| INV_RAT | | 1.719 ** (0.854) | 1.556 * (0.869) | -1.615 (1.605) | |
| VAL_RAT | | | -0.627 ** (0.312) | -0.644 ** (0.313) | -0.005 |
| INV_RAT*EQ_RAT | | | | -5.157 *** (2.012) | -0.044 |
| | | | | | |
| McFadden R squared | 0.213 | 0.218 | 0.223 | 0.231 | |

Equation:

$$PD_i = P(Y_{i,t+1}=1) = \frac{\exp(\alpha + \sum_{k=1}^{R} \beta_k X_{ik,t})}{1+\exp(\alpha + \sum_{k=1}^{R} \beta_k X_{ik,t})} \qquad Y_{i,t+1 \quad i=1,\ldots,n} = \begin{cases} 1 \; if \;\; obligor \; i \; defaults \; in \; t+1 \\ 0 \; if \;\; obligor \; i \; does \; not \; defaults \; in \; t+1 \end{cases} \qquad X_{ik,t \quad i=1,\ldots,n} = k^{th} \; regressor \;\; for \; obligor \; i \; in \; t$$

Notes: Each regression includes the dummies listed in Table 1; * significant at 10%; ** significant at 5%; *** significant at 1%; SE in parenthesis. The Hosmer-Lemenshow test supports the estimated models.

[11] In order to perform out-of-sample analysis (see Section 5.1), the sample has been divided so that the estimation is performed over 2/3 of the full dataset while the remaining 1/3 is left for out-of-sample checks. See Stein (2002) for the selection of the out-of-sample dataset.

Then, to test whether the dimension of the patent portfolio reduces the innovative firm's riskiness, we add to the balance-sheet regressors the variable that reflects the R&D productivity, discussed in Section 4 and recalled in Table 1 (INV_RAT). The estimation of Model 2 confirms the results obtained in the baseline model with balance-sheet variables only, but it highlights an interesting result concerning the patent productivity coefficient which is positive and significant. In other words, for a given level of R&D expenses, a larger patent portfolio implies a higher firm's riskiness. This finding is consistent with Motohashi (2011), that stresses the role of risk and uncertainty on the good fate of innovation activities and with the literature on innovation. The latter in fact emphasizes the commercialization risk of a given technology in particular for small companies, which typically do not control the complementary assets required for a successful exploitation of that technology (e.g. Arora et al. 2001).

The question at this stage is to understand how this result can be explained. To this end we assess the natural role played by the value, beside the dimension, of the patent portfolio and we add the second innovation-related regressor: VAL_RAT, discussed in Section 4.1 and recalled in Table 1. By inspection of Model 3, the two different facets of innovation clearly emerge to act in opposite directions: while the value of the patent portfolio contributes to decrease the PD, its dimension appears to be signaling a higher firm's riskiness. Comparing Model 2 and 3, it is apparent that, when the value of the portfolio is taken into account, the importance of the R&D productivity decreases also in term of significance. However, the portfolio value alone does not cancel the role played by the portfolio productivity, i.e. the riskiness connected with the portfolio dimension.

In order to solve for this apparent puzzle, in Model 4 we further investigate the relevance of this variable in connection with the capital structure of the firm. To this end, we interact the patent productivity with the variable representing the firm's leverage, i.e. the equity over debt ratio (EQ_RAT). This model shows that, while the patent value remains significant, the patent productivity alone loses explanatory power, but when it is interacted with the EQ_RAT it turns to be a very strong predictor of the default event with an elasticity of 4.4%.

In sum, as for the innovation-related variable, by comparative inspection of Model 3 and 4, our results show that the patent value always reduces the PD as expected, while the patent number per se does reduce the PD only if supported by an appropriate equity level. [12]

[12] Overall, the size of the balance-sheet variables' coefficients is stable and robust across all specifications: the variable with highest impact on the probability of default event is liquidity (LIQ_RAT) followed by the two profitability ones (PROF_RAT and EX_RAT). In particular, one standard deviation increase in the liquidity reduces the PD by 2.8% whereas in terms of profitability by 1.4%. We find that the PD moderately decreases with firm size as approximated by sales.

This finding is consistent with the theory on the existence of financial constraints for innovative SMEs discussed previously. On the one hand, SMEs use more equity to finance the innovation activities because of the presence of the investor's information asymmetries on the quality of their assets. On the other, because more innovative SMEs face relatively tighter adverse selection in the credit market, only a few of them, which have appropriate equity, will develop those assets. Hence the real effect of financial constraints may be plausibly even more severe than the evidence suggested by results in Table 2.

The overall goodness of fit of Model 4 is more than 23.1%, which is not small given the limited number of variables of the model and the fact that our sample is made up of the cross-section dimension only. The increase in the explanatory power of Model 4 with respect to Model 1 is 8.45% and is due to the inclusion of the innovative variables and the term accounting for the interaction between innovation and firm's equity.

**5.1 Additional analyses**

In order to further assess the validity of the model, we perform additional goodness of fit tests. The first one is the Cumulative Accuracy Profile (CAP) that measures simultaneously Type I and Type II errors. In the CAP analysis companies are ranked by fitted values of the PD event. For a given percentage of the observations x, a CAP curve is built by computing the percentage of actual default events with the risk score equal to or lower than x (for a more detailed illustration see e.g. Sobehart et al., 2001).

In Figure 1 the thick curve shows the goodness of the estimated model. It depicts the percentage of actual default events (vertical axis) versus the defaults predicted by the model (horizontal axis). The diagonal line represents the case of non-informative model, whereas the upper line the perfectly predicting model. In our case the model shows a high predictive power estimating about 50% of the defaulters within only 6% of the observation. A more synthetic measure is the Accuracy Ratio (AR) which graphically equals the area predicted by the CAP divided by the area of the perfectly predicting model: in-sample it is 70.2% (see Figure 1). The model performs well also out-of-sample: the CAP out-of sample follows closely the dynamics of that in-sample though the Accuracy Ratio is smaller (Figure 2).

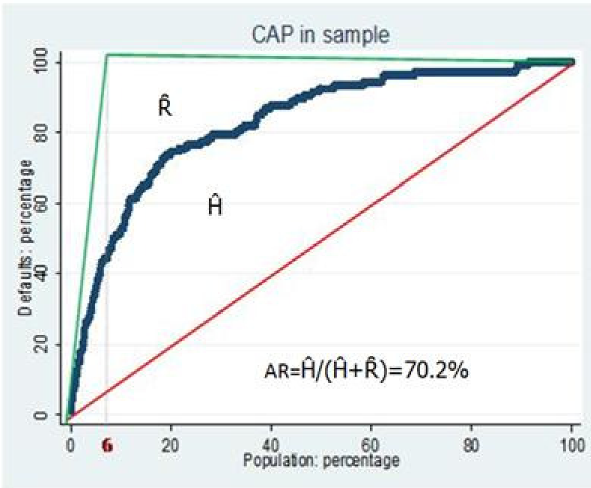


Figure 1 In-sample Cumulative Accuracy Profile and Accuracy Ratio

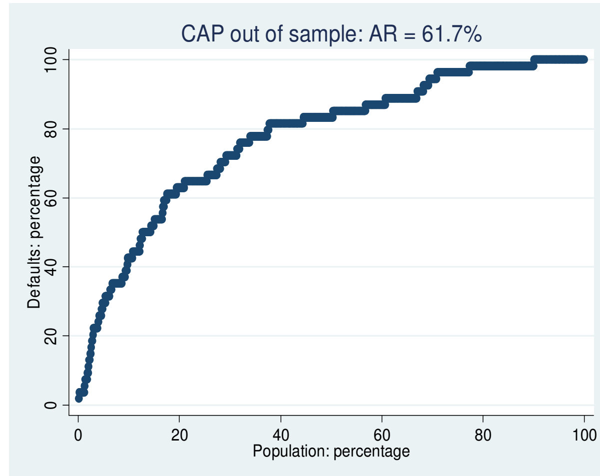


Figure 2 Out-of-sample Cumulative Accuracy Profile and Accuracy Ratio

Moreover, the two types of error of the in-sample and out-of-sample dataset closely co-evolve, which again strongly support the validity of the model estimated in Table 2 (Figure 3).

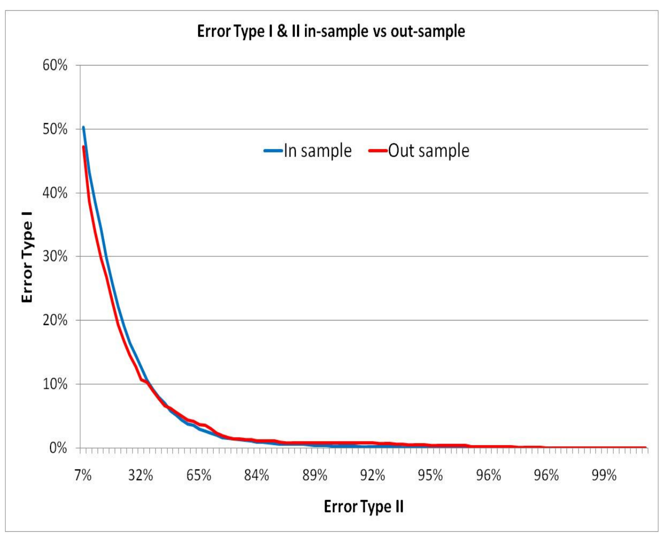


Figure 3 Comparison of Prediction Errors

Figure 4a and 4b compare the CAP curves for Model 1 and Model 4 in- and out-of sample respectively. The model with innovative variables shows a higher explanatory power for the medium risk firms (percentiles 15-35 percentiles for the in-sample dataset), whereas for the high risk firms (top 10% of the distribution) the models works similarly. In other words, while financial variables are quite informative for the most risky firms, innovative variables contribute to add information when the credit quality of the enterprise is less clear-cut and hence more difficult to gauge.

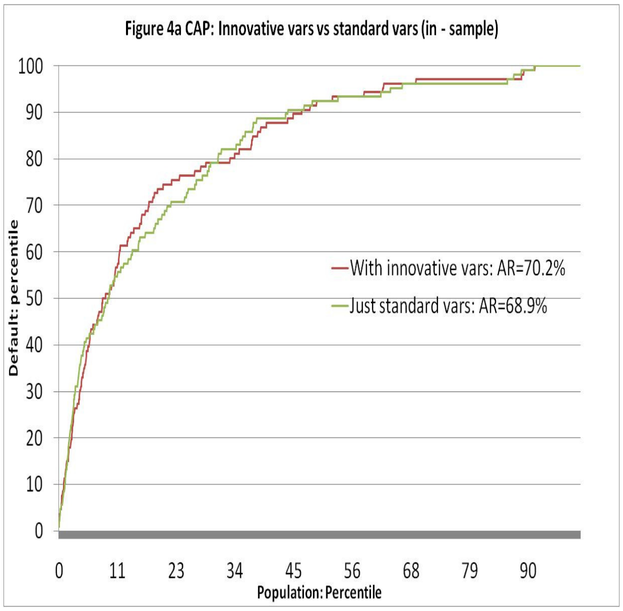


Figure 4a Comparison of Model 1-4 in sample

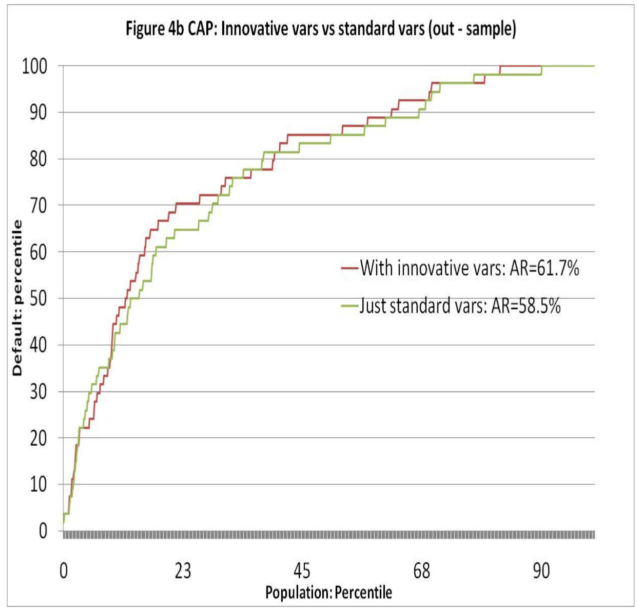


Figure 4b Comparison of Model 1-4 out sample

In order to gauge the increased accuracy obtained by the inclusion of the patent-related variable, we now directly compare the baseline model (Model 1) with the one proposed in this paper (Model 4).

Table 4. A comparison of accuracy ratios

| *Accuracy Ratio (AR)* | *Model 1* | *Model 4* |
|---|---|---|
| AR in-sample | 0.689 | 0.702 |
| AR out-of-sample | 0.585 | 0.617 |

In terms of Accuracy Ratio, Table 4 highlights that the model we propose performs better both in-sample and out-of-sample. Moreover, the relative reduction of the AR in the out-of-sample dataset for the models with innovative variables is lower (12.1%) than in the case of the Model 1 (15.1%) thus indicating that our model is more accurate in default prediction.

## 6. Conclusions

In this paper we develop a parsimonious logit regression model to estimate a PD of innovative SMEs in EU15 countries with default years 2006-2008. To the best of our knowledge this is among the first attempts which combines accounting and innovation-related variables to predict the default event. Based on the signaling value of patents, we include their consideration in an econometric model for the PD prediction of innovative SMEs, which builds on one of the most widespread model based on accounting data. To this end, we must also tackle the issue of defining and measuring the patent-related regressors.

We first test a standard specification of the model relying on accounting variables only, which gives results in line with the literature (e.g. Altman and Sabato, 2007). Second, by including a regressor that captures the innovation productivity, we show that the (normalized) dimension of the patent portfolio increases the innovative firm's PD. This result loses some strength but remains valid even after accounting for the patents' value, which points in the opposite direction contributing to a reduction in the PD. Given that the value of the patent portfolio is not enough to explain the riskiness related to its dimension, we address a final question concerning the role played by the capital structure of the firm in connection with the innovation activity and we introduce a term that interacts innovation productivity and the equity ratio.

In sum, our results show that, while the value of the patent portfolio always reduces the PD, its dimension increases the firm's riskiness unless it is coupled with an appropriate equity level. Moreover the model proves to have a higher in- and out-of sample accuracy if compared with the standard model based on accounting variables only.

The model proposed in this paper to predict the PD of innovative firms can help in reducing the asymmetric information issues which are particularly pronounced for these enterprises. It can thus be useful for banks and investors interested in gauging the riskiness of this type of firms in consideration of their peculiar features, which relates to their innovative value and potential.

Concluding, it is noteworthy to recall that patents are not the only information trail to reveal the technological and commercial potential of a start-up. Other studies have claimed that web newswires could constitute an additional information sources for financial investors (see Kerr et al., 2010). We aim to develop similar measures and use them for default prediction in our future research-work.

**Acknowledgements**

We would like to thank for comments and suggestions Kazu Motohashi, Elisabeth Müller, Raffaele Oriani, Francesco Pattarin, Georg Licht, and participants at the CEFIN Workshop (Modena, 2010), 4th CSDA International Conference on Computational and Financial Econometrics (CFE'10, London), 4th ZEW Conference on Economics of Innovation and Patenting (2011, Mannheim), 3rd Workshop on The Output of R&D Activities: Harnessing the Power of Patents Data (2011, Sevilla), International Risk Management Conference (2011, Amsterdam), 15th International Conference on Insurance: mathematics and Economics (2011, Trieste), IFABS Conference (2011, Rome). Usual caveat apply. Chiara Pederzoli and Costanza Torricelli also acknowledge financial support from MIUR-PRIN 2007.

**References**

Altman, E.I., 1968. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy", Journal of Finance, 13, 589–609.

Altman, E.I., Hotchkiss, E., 2006. Corporate Financial Distress and Bankruptcy: Predict and Avoid Bankruptcy, Analyse and Invest in Distressed Debt, 3rd Edition, Hardcover: Wiley Finance.

Altman, E.I., Sabato, G., 2005. Effects of the New Basel Capital Accord on Bank Capital Requirements for SMEs, Journal of Financial Services Research, 28, 15–42.

Altman, E.I., Sabato, G., 2007. Modelling Credit Risk for SMEs: Evidence from the U.S. Market, Abacus, 43, 3, 332-357.

Arora A, Fosfuri A,  Gambardella A., 2001. Markets for Technology: The Economics of Innovation and Corporate Strategy. Cambridge, MA: MIT Press.

Arrow, K.J., 1975. Vertical Integration and Communication, Bell Journal of Economics 6, 173-183.

Basel Committee on Banking Supervision (BCBS), 2006. International Convergence of Capital Measurement and Capital Standards: A Revised Framework - Comprehensive Version, Bank for International Settlements.

Beaver, V., 1966. Financial Ratios as Predictors of Failure, Journal of Accounting Research, 5, 71-111.

Behr, P., Güttler, A., 2007. Credit Risk Assessment and Relationship Lending: an Empirical Analysis of German Small and Medium-sized Enterprises, Journal of Small Business Management, 45, 2, 194–213.

Brown, J. R., Fazzari, S. M., Petersen, B. C., 2009. Financing Innovation and Growth: Cash Flow, External Equity and the 1990s R&D Boom, Journal of Finance, 64, 1, 151-185.

Buddelmeyer, H., Jensen, P., Wesbster, E., 2010. Innovation and the determinants of company survival, Oxford Economic Papers, 62, 2, 261-285.

Carpenter, R. E., Petersen, B. C., 2002. Capital Market Imperfections, High-Tech Investment, and New Equity Financing, The Economic Journal, 112, 54-72.

Denis, D. 2004. Entrepreneurial Finance: an Overview of the Issues and Evidence, Journal of Corporate Finance, 10, 301– 326.

EIB, 2009. R&D and the Financing of Innovation in Europe, European Investment Bank vol 14, 2.

Fantazzini D., Figini S., 2009. Default Forecasting for Small-medium Enterprises: Does Heterogeneity Matters?, International Journal of Risk Assessment and Management, 11, 1-2, 38-49.

Fidrmuc J., Heinz C., 2009. Default Rates in the Loan Market for SMEs: Evidence from Slovakia, University of Munich, IFO Working Paper N. 72.

Gambardella, A., Harhoff, D., Verspagen, B., 2008. The Value of European Patents. European Management Review, 5, 2, 69-84.

Griliches, Z., 1990. Patent Statistics as Economic Indicators: A Survey. Journal of Economic Literature XXVIII, December, 1661-1707.

Grunert, J., Norden, L., Weber, M., 2005. The Role of Non-financial Factors in Internal Credit Ratings, Journal of Banking & Finance, 29, 2, 509–531.

Guellec, D., Van Pottelsberghe de la Potterie, B., 2000. Applications, Grants, and the Value of Patent, Economic Letters 69, 1, 109-114.

Guellec, D., van Pottelsberghe de la Potterie, B., 2007. The Economics of the European Patent System: IP Policy for Innovation and Competition, Oxford University Press, New York.

Hall, B.H., 2002. The Financing of Research and Development, NBER Working Paper No. 8773 February 2002.

Hall, B. H., 2005. Measuring the Returns to R&D: The Depreciation Problem, Annales d'Economie et de Statistique 79/80, 341-381.

Hall, B. H., Jaffe, A. B., Trajtenberg, M., 2001. The NBER Patent Citations Data File: Lessons, Insights, and Methodological Tools. Cambridge, MA: NBER Working Paper No. 7741.

Hall, B. H., Jaffe, A. B., Trajtenberg, M., 2005. Market Value and Patent Citations. RAND Journal of Economics 36, 16-38.

Hall B. H. and Lerner, J., 2010. Financing R&D and Innovation, in the Handbook of the Economics of Innovation, B. H. Hall and N. Rosenberg (eds.), April 2010, Elsevier.

Hall, B.H., Mairesse, J., Mohnen, P., 2010. Using Innovation Surveys for Econometric Analysis, in the Handbook of the Economics of Innovation, B. H. Hall and N. Rosenberg (eds.), April 2010, Elsevier.

Hall, B. H., Thoma, G., Torrisi, S., 2007. The Market Value of Patents and R&D: Evidence from European firms, NBER Working paper No. 13426, Cambridge, Mass.

Harhoff, D., 1996. Strategic Spillovers and Incentives for Research and Development, Management Science, 42, 6, 907-925.

Harhoff, D., Körting, T., 1998. Lending Relationships in Germany: Empirical Results from Survey Data, Journal of Banking and Finance, 22, 1317-1353.

Harhoff, D., Scherer, F. M., Vopel, K., 2003. Citations, Family Size, Opposition and the Value of Patent Rights - Evidence from Germany, Research Policy, 32, 1343-1363.

Harhoff D., Häussler, C., Müller. E., 2009. To be Financed or Not: The Role of Patents for VC Financing, CEPR Discussion Paper 7115, 2009.

Harhoff, D., Thoma, G., 2010. Inventor Location and the Globalization of R&D, Paper presented at the Conference on Patent Statistics on Decision Makers, 17-18 November 2010, Vienna, Austria.

Hsu, D., Ziedonis R., 2008. Patents as Quality Signals for Entrepreneurial Ventures, Academy of Management Best Paper Procedings, Anaheim California, 2008.

Kerr, W. R., Lerner, J., Schoar, A., 2010. The Consequences of Entrepreneurial Finance: A Regression Discontinuity Analysis, NBER Working Paper No. 15831, Cambridge, Mass.

Lanjouw, J. O., Schankerman, M., 2004. Patent Quality and Research Productivity: Measuring Innovation with Multiple Indicators. Economic Journal, 114, 441-465.

Lennox, C., 1999. Identifying Failing Companies: A Reevaluation of the Logit, Probit and DA Approaches, Journal of Economics and Business, 51, 347–364.

Lerner, J., 1994. The Importance of Patent Scope: An Empirical Analysis, Rand Journal of Economics 25, 2, 319-33.

Motohashi, K., 2011. Innovation and Entrepreneurship: A first look at linkage data of Japanese patent and enterprise census, RIETI Discussion Paper Series 11-E-007, February.

OECD, 2009a. OECD Patent Statistics Manual, OECD, Paris.

OECD, 2009b. Financing High Growth and Innovative Start-ups and SMEs: Data and Measurement Issues, OECD, Paris.

Ohlson, J., 1980. Financial Ratios and the Probabilistic Prediction of Bankruptcy, Journal of Accounting Research, 18, 109-131.

Pederzoli, C., Torricelli, C., 2010. A Parsimonious Default Prediction Model for Italian SMEs, Banks and Bank Systems, 5, Issue 4, 28-32

Sobehart, J., Keenan, S. Stein, R., 2001. Benchmarking Quantitative Default Risk Models: A Validation Methodology, Algo Research Quarterly, 4, Nos.1/2 March/June 2001, 57-72.

Stein, R. M., 2007. Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation, Journal of Risk Model Validation, 1, 1, 77-113.

Thoma, G., Torrisi, S., Gambardella, A., Guellec D., Hall, B. H., Harhoff, D., 2010. Harmonizing and Combining Large Datasets – An Application to Firm-Level Patent and Accounting Data, NBER Working Paper No. 15851, March 2010.

Trajtenberg, M., 1990. A Penny for Your Quotes: Patent Citations and the Value of Innovations. Rand Journal of Economics 21, 172-187.

**Appendix A – The sample: descriptive statistics**

Table A1. Firms in the sample

| | Cohort | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Industries | *pre-1970* | *1971-80* | *1981-90* | *1991-95* | *1996-2000* | *post-2000* | *Overall* | *%* |
| | | | | | | | | |
| *Chem.&Phar.* | 62 | 27 | 39 | 11 | 17 | 3 | 159 | 6.0% |
| *Low tech* | 275 | 140 | 218 | 77 | 88 | 15 | 813 | 30.5% |
| *Manufact.* | 268 | 175 | 231 | 120 | 122 | 36 | 952 | 35.7% |
| *Retail Distr.* | 73 | 54 | 113 | 70 | 50 | 17 | 377 | 14.1% |
| *Services* | 35 | 27 | 91 | 75 | 93 | 43 | 364 | 13.7% |
| | | | | | | | | |
| *Overall* | 713 | 423 | 692 | 353 | 370 | 114 | 2,665 | 100% |
| *Overall %* | 26.8% | 15.9% | 26.0% | 13.2% | 13.9% | 4.3% | 100% | |

Table A2. Defaulted firms in the sample

| | Cohort | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Industries | *pre-1970* | *1971-1980* | *1981-1990* | *1991-1995* | *1996-2000* | *post-2000* | *Overall* | *%* |
| *Chem. & Phar.* | 2 | 3 | 5 | 1 | 2 | 0 | 13 | 8.1% |
| *Low tech* | 11 | 13 | 18 | 3 | 3 | 2 | 50 | 31.3% |
| *Manufacturing* | 16 | 9 | 15 | 12 | 9 | 2 | 63 | 39.4% |
| *Retail Distrib.* | 5 | 2 | 5 | 5 | 6 | 0 | 23 | 14.4% |
| *Services* | 0 | 0 | 3 | 1 | 5 | 2 | 11 | 6.9% |
| | | | | | | | | |
| *Overall* | 34 | 27 | 46 | 22 | 25 | 6 | 160 | 100% |
| *Overall %* | 21.3% | 16.9% | 28.8% | 13.8% | 15.6% | 3.8% | 6.0% | |

Note: Low tech industries include US SIC code 10-33, such as agriculture, forestry, fishing, mining, food and wood products, and textiles, but not chemicals and pharmaceuticals.

**Appendix B – The variables: descriptive statistics**

| *All sample* | | | | | | *Chemicals and Pharmaceuticals* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *variable* | *mean* | *sd* | *median* | *1Q* | *3Q* | *variable* | *mean* | *sd* | *median* | *1Q* | *3Q* |
| EQ_RAT | 0.824 | 1.481 | 0.443 | 0.171 | 0.958 | EQ_RAT | 1.030 | 1.274 | 0.589 | 0.206 | 1.258 |
| LIQ_RAT | 0.220 | 1.802 | 0.034 | 0.007 | 0.106 | LIQ_RAT | 0.212 | 1.114 | 0.031 | 0.005 | 0.118 |
| PROF_RAT | 0.002 | 0.288 | 0.022 | -0.005 | 0.072 | PROF_RAT | 0.014 | 0.136 | 0.026 | -0.012 | 0.072 |
| | - | | | | | | - | | | | |
| EX_RAT | 0.003 | 0.111 | 0.000 | 0.000 | 0.001 | EX_RAT | 0.008 | 0.045 | 0.000 | -0.003 | 0.002 |
| COV_RAT | 0.290 | 92.289 | 4.000 | 0.475 | 19.687 | COV_RAT | 0.283 | 111.133 | 3.999 | -0.503 | 19.244 |
| SALES | 2.785 | 1.815 | 2.569 | 1.366 | 3.900 | SALES | 3.478 | 1.770 | 3.569 | 2.124 | 4.736 |
| | - | | | | | | - | | | | |
| VAL_RAT | 0.383 | 0.385 | -0.426 | -0.676 | -0.106 | VAL_RAT | 0.217 | 0.401 | -0.206 | -0.506 | 0.002 |
| INV_RAT | 0.154 | 0.108 | 0.136 | 0.079 | 0.229 | INV_RAT | 0.116 | 0.092 | 0.097 | 0.056 | 0.146 |

| *Low tech industries* | | | | | | *Manufacturing* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *variable* | *mean* | *sd* | *median* | *1Q* | *3Q* | *variable* | *mean* | *sd* | *median* | *1Q* | *3Q* |
| EQ_RAT | 0.802 | 1.277 | 0.459 | 0.209 | 0.963 | EQ_RAT | 0.687 | 1.088 | 0.428 | 0.190 | 0.829 |
| LIQ_RAT | 0.081 | 0.188 | 0.029 | 0.006 | 0.079 | LIQ_RAT | 0.111 | 0.467 | 0.035 | 0.006 | 0.094 |
| PROF_RAT | 0.018 | 0.158 | 0.022 | -0.002 | 0.071 | PROF_RAT | 0.009 | 0.262 | 0.024 | -0.006 | 0.070 |
| | | | | | | | - | | | | |
| EX_RAT | 0.003 | 0.084 | 0.000 | -0.001 | 0.001 | EX_RAT | 0.001 | 0.064 | 0.000 | 0.000 | 0.001 |
| COV_RAT | 0.260 | 80.203 | 3.875 | 0.780 | 14.806 | COV_RAT | 0.286 | 86.507 | 4.274 | 0.704 | 19.476 |
| SALES | 3.076 | 1.794 | 2.928 | 1.752 | 4.022 | SALES | 2.963 | 1.762 | 2.768 | 1.713 | 3.955 |
| | - | | | | | | - | | | | |
| VAL_RAT | 0.384 | 0.419 | -0.444 | -0.701 | -0.120 | VAL_RAT | 0.430 | 0.332 | -0.469 | -0.682 | -0.190 |
| INV_RAT | 0.164 | 0.111 | 0.161 | 0.092 | 0.229 | INV_RAT | 0.160 | 0.114 | 0.142 | 0.083 | 0.229 |

| *Retail Distribution* | | | | | | *Services* | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| *variable* | *mean* | *sd* | *median* | *1Q* | *3Q* | *variable* | *mean* | *sd* | *median* | *1Q* | *3Q* |
| EQ_RAT | 0.992 | 2.383 | 0.373 | 0.129 | 0.923 | EQ_RAT | 0.986 | 1.647 | 0.470 | 0.138 | 1.136 |
| LIQ_RAT | 0.346 | 2.798 | 0.030 | 0.008 | 0.098 | LIQ_RAT | 0.718 | 3.847 | 0.077 | 0.018 | 0.312 |
| | | | | | | | - | | | | |
| PROF_RAT | 0.001 | 0.254 | 0.019 | -0.001 | 0.069 | PROF_RAT | 0.065 | 0.561 | 0.017 | -0.053 | 0.084 |

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | - | | | | | | - | | | | |
| EX_RAT | 0.003 | 0.062 | 0.000 | -0.001 | 0.000 | EX_RAT | 0.019 | 0.255 | 0.000 | 0.000 | 0.000 |
| COV_RAT | 0.374 | 104.939 | 3.559 | 1.000 | 25.875 | COV_RAT | 0.270 | 107.255 | 3.667 | -3.510 | 30.464 |
| SALES | 2.225 | 1.643 | 2.114 | 0.832 | 3.252 | SALES | 1.889 | 1.765 | 1.294 | 0.382 | 3.004 |
| | - | | | | | | - | | | | |
| VAL_RAT | 0.402 | 0.399 | -0.456 | -0.718 | -0.044 | VAL_RAT | 0.313 | 0.384 | -0.339 | -0.594 | -0.001 |
| INV_RAT | 0.148 | 0.094 | 0.141 | 0.075 | 0.229 | INV_RAT | 0.142 | 0.097 | 0.117 | 0.074 | 0.229 |

# Previously published “CEFIN Working Papers”

24 *Market Reaction to Second-Hand News: Attention Grabbing or Information Dissemination?*, by Cervellati E.M., Ferretti R., Pattitoni P. (January 2011)

23 *Towards a volatility index for the Italian stock market,* by Muzzioli S. (September 2010)

22 *A parsimonious default prediction model for Italian SMEs*, by Pederzoli C., Torricelli C. (June 2010)

21 *Average Internal Rate of Return and investment decisions: a new perspective,* by Magni C.A. (February 2010)

20 *The skew pattern of implied volatility in the DAX index options market,* by Muzzioli S. (December 2009)

19 *Accounting and economic measures: An integrated theory of capital budgeting,* by Magni C.A. (December 2009)

18 *Exclusions of US-holders in cross-border takeover bids and the principle of equality in tender offers*, by Mucciarelli F. (May 2009).

17 *Models for household portfolios and life-cycle allocations in the presence of labour income and longevity risk*, by Torricelli C. (March 2009)

16 *Differential evolution of combinatorial search for constrained index tracking*, by Paterlini S, Krink T, Mittnik S. (March 2009)

15 *Optimization heuristics for determining internal rating grading scales*, by Paterlini S, Lyraa M, Pahaa J, Winker P. (March 2009)

14 *The impact of bank concentration on financial distress: the case of the European banking system*, by Fiordelisi F, Cipollini A. (February 2009)

13 *Financial crisis and new dimensions of liquidity risk: rethinking prudential regulation and supervision*, by Landi A, Gualandri E, Venturelli V. (January 2009)

12 *Lending interest rate pass-through in the euro area: a data-driven tale*, by Marotta G. (October 2008)

11 *Option based forecast of volatility: an empirical study in the Dax index options market*, Muzzioli S. (May 2008)

10 *Lending interest rate pass-through in the euro area*, by Marotta G. (March 2008)

9 *Indebtedness, macroeconomic conditions and banks’ losses: evidence from Italy*, by Torricelli C, Castellani S, Pederzoli C. (January 2008)

8 *Is public information really public? The role of newspapers*, Ferretti R, Pattarin F. (January 2008)

7 *Differential evolution of multi-objective portfolio optimization*, by Paterlini S, Krink T. (January 2008)

6 *Assessing and measuring the equity gap and the equity*, by Gualandri E, Venturelli V. (January 2008)

5 *Model risk e tecniche per il controllo dei market parameter*, Torricelli C, Bonollo M, Morandi D, Pederzoli C. (October 2007)

4 *The relations between implied and realised volatility, are call options more informative than put options? Evidence from the Dax index options market*, by Muzzioli S. (October 2007)

3 *The maximum LG-likelihood method: an application to extreme quantile estimation in finance*, by Ferrari D., Paterlini S. (June 2007)

2 *Default risk: Poisson mixture and the business cycle*, by Pederzoli C. (June 2007)

1 *Population ageing, household portfolios and financial asset returns: a survey of the literature*, by Brunetti M. (May 2007)