The Journal of Credit Risk (1–33) Volume 6/Number 2, Summer 2010

The value of non-financial information in small and medium-sized enterprise risk management

Edward I. Altman
NYU Salomon Center, Leonard N. Stern School of Business,
New York University, 44 West 4th Street, New York, NY 10012, USA;
email: ealtman@stern.nyu.edu

Gabriele Sabato
Group Credit Risk, Royal Bank of Scotland, Gustav Mahlerlaan 10,
1000EA Amsterdam, The Netherlands; email: gasabato@hotmail.com

Nicholas Wilson
Credit Management Research Centre, Leeds University Business School,
Leeds LS2 9JT, UK; email: nw@lubs.leeds.ac.uk

Few studies that have focused on developing credit risk models specifically for small and medium-sized enterprises (SMEs) have included non-financial information as a predictor of company creditworthiness. In this study we have available non-financial, regulatory compliance and "event" data to supplement the limited accounting data that is often available for non-listed firms. We employ a sample consisting of over 5.8 million sets of accounts of unlisted firms, of which over 66,000 failed during the period 2000–2007. We find that data relating to legal action by creditors to recover unpaid debts, company filing histories, comprehensive audit report/opinion data and firm-specific characteristics make a significant contribution to increasing the default prediction power of risk models built specifically for SMEs.

1 INTRODUCTION

The Basel Capital Accord and the 2007 financial crisis have provided renewed impetus for lenders to research and develop adequate failure prediction models for all of the corporate and retail sectors of their lending portfolios. The Basel II definition of financial distress, 90 days overdue on credit agreement payments, is the operational

The authors are grateful to participants at the IRMC Conference in Florence (June 2008) and to the editor and reviewers of this journal for very helpful comments.

definition for major lenders. The literature on the modeling of credit risk for large, listed companies is extensive and gravitates toward either of two approaches: the z-score approach of using historical accounting data to predict insolvency (see, for example, Altman (1968)); and models that rely on securities market information (Merton (1974)). In retail lending, risk modeling can be undertaken using very large samples of high-frequency consumer data and combinations of in-house portfolio data (eg, payment history) and bureau data from the credit reference agencies to develop proprietary models.

Prior to the introduction of Basel II, retail lending was mainly synonymous with consumer lending. Since the accord an increasing number of banks have started to reclassify commercial clients from the corporate area into the retail area. Although this decision may have originally been motivated by expected capital savings (see Altman and Sabato (2005)), financial institutions have now realized that the major benefits are on the efficiency and profitability side. Banks are also realizing that small and medium-sized enterprises (SMEs) are a distinct kind of client with specific needs and peculiarities that require risk-management tools and methodologies specifically developed for them (see Altman and Sabato (2007)).

Indeed, SMEs are the predominant type of business in all Organisation for Economic Cooperation and Development economies and typically account for two-thirds of all employment. In the UK, unlisted firms make up the majority of firms that ultimately fail. Of the 1.2 million active companies that are registered less than 12,000 are listed on the stock market. In the US, private companies contribute over 50% of GDP (Shanker and Astrachan (2004)). The flow of finance to this sector is much researched as it is seen to be crucial to economic growth and success. However, from a lending perspective, research on credit risk management for small companies is relatively scarce. The best way of ensuring a flow of finance to SMEs is to improve credit information and to develop adequate risk models for this sector.

Techniques for modeling corporate insolvency have long been applied as a means of assessing and quantifying the risk of listed companies, and the research into failure-rate prediction has focused almost exclusively on listed companies. Much of the pioneering work on bankruptcy prediction has been undertaken by Altman (1968, 1993).[1] These earlier works were undertaken primarily during the 1960s, although extensions of this work to developing countries appeared during the 1990s (see Altman and Narayanan (1997)). Early research into corporate failure prediction involved determining which accounting ratios best predict failure, primarily employing multiple discriminant analysis (MDA) or logit/probit models. In most of these accounting-ratio-based studies, ratios are calculated at a predetermined time before bankruptcy

[1] Altman (1993) is the second edition of a book first published in 1983. A third edition, co-authored by Edith Hotchkiss, was published in 2005.

(usually one year) and as such these models are often referred to as static models. Notably, these studies focus on the use of data other than accounting data, for example, von Stein and Ziegler (1984) examine the impact of managerial behavior on failure. This earlier work invariably suffers from having only a small sample of failed firms available for analysis.

Recently, Altman and Sabato (2007) applied, with some success, a distress prediction model estimated specifically for the US SME sector based on a set of financial ratios derived from accounting data. They demonstrate that banks should not only apply different procedures (in the application and behavioral process) to manage SMEs as compared with large corporate firms, but these organizations should also use scoring and rating systems specifically addressed to the SME portfolio. The lack of any non-financial and compliance information about the companies in the sample is a significant limiting factor, forcing them to exclude a relevant portion of small companies without accounting data.

In practice, the building of credit risk models for private companies is necessarily limited by data availability. Of course, market data is not available for unlisted firms. Furthermore, many unlisted firms are granted concessions regarding the amount of financial statement data they are required to file, meaning that data required to calculate some of the accounting ratios employed in studies of the failure of listed companies is not available for SMEs. In recognition of the paucity of data available for many non-listed firms, a paper by Hol (2007) analyzes the incremental benefit of employing macroeconomic data to predict bankruptcy on a sample of Norwegian unlisted firms. Other studies focus on specifying alternative outcome definitions. Peel and Peel (1989) use a multi-logit approach to modeling financial distress in preference to the usual binary outcome. Peel and Wilson (1989) estimate a multi-logit model that identifies “distressed acquisitions” as an important outcome from bankruptcy situations. Fantazzini and Figini (2008) propose a non-parametric approach based on a random survival forest and compare its performance with a standard logit model.

Recent literature has highlighted the benefits of including variables such as age and type of business, industrial sector, etc, in combination with financial ratios (Grunet *et al* (2004)). Peel and Peel (1989) and Peel *et al* (1986) provide evidence from a UK sample that the timeliness of financial reporting is a potential indicator of financial stress. However, these studies do not focus on SME clients and a very limited amount of non-financial information is analyzed and used for modeling purposes.

In this study we update the current literature in several ways. First, we test the Altman and Sabato (2007) SME model on a geographically different sample (UK companies) including an extremely high number of small companies (5.8 million) covering a very recent economic period (2000–2007). In doing so we eventually prove the substantial soundness and significant prediction power of our SME default prediction model. Second, by using a unique data set[2] we are able to explore the value added by non-financial and compliance information specifically for SMEs. We find that this information, when available, is likely to significantly improve the prediction accuracy of the model. Last, using the available non-financial information we also develop a default prediction model for the large number of SMEs for which financial information is very limited (eg, sole traders, professionals, micro-companies, companies that choose simplified accountancy or tax reporting). Solutions that address credit risk management for clients such as these have never been provided by the existing literature.

The database available covers the UK companies that filed accounts in the period 2000–2007. The data consists of over 5.8 million records of accounting and other publicly available data for companies active in this period. The incidence of insolvency in the data covers 66,833 companies (1.2% of the total). Moreover, a subset of SMEs based in the UK have account-filing exemptions, which means that the amount of accounting data available for these companies is quite limited. These companies usually represent more than 60% of the total number of SMEs. Thus, small-company[3] accounts include an abbreviated balance sheet and no profit-and-loss account and medium-sized-company[4] accounts include a full balance sheet but an abbreviated profit-and-loss account. We have access to some profit-and-loss account data for around 40% of our unlisted companies.

The remainder of the paper proceeds as follows. In Section 2 we provide an overview of the definition of SMEs, of the definition of failure and of the extant literature on failure prediction. In Section 3 we provide details of our UK sample and we undertake a detailed examination of the data available to us to predict small-business failure among unlisted firms. In Section 4 we present a failure prediction model for SMEs for which profit-and-loss data is available. In particular, we are able to estimate the Altman

---

[2] We are grateful to members of the Credit Management Research Centre (Leeds University) and CreditScorer Ltd for the work undertaken in compiling this data set, particularly to Ali Altanlar for work on the data set and to Paula Hill for undertaking exploratory work on an earlier subsample of the data.

[3] UK companies are required to file accounts at Companies House. The Companies House website (URL: www.companieshouse.gov.uk) defines a small company as one for which at least two of the following conditions are met: annual turnover is £5.6 million or less; the balance sheet total is £2.8 million or less; and the average number of employees is 50 or fewer. For financial years ending before January 30, 2004 the exemptions thresholds are £2.8 million for turnover and £1.4 million for the balance sheet total.

[4] The Companies House website defines a medium-sized company as one for which at least two of the following conditions are met: annual turnover is £22.8 million or less; the balance sheet total is £1.4 million or less; and the average number of employees is 250 or fewer. For financial years ending before January 30, 2004 the exemptions thresholds are £11.2 million for turnover and £5.6 million for the balance sheet total.

and Sabato model on the UK sample (SME1). We test for the impact of adding non-financial and "event" data to the models estimated in Section 4. In addition we present a failure prediction model for small firms that do not file profit-and-loss statement data (SME2). In Section 5 we employ our models to undertake out-of-sample forecasts. Section 6 provides a concluding discussion.

2 REVIEW OF THE RELEVANT RESEARCH LITERATURE ON SMALL AND MEDIUM-SIZED ENTERPRISES

Small-business lending has mainly received attention from researchers and practitioners in the last 10 years. Several aspects have been analyzed. From these studies the issues that are most relevant to this study are the definition of SMEs, the definition of SME failure and the modeling techniques that can be applied to predict small-business failure. This section reviews some of the most important work on these subjects.

2.1 SME definition

It seems that there is no common definition of SMEs across different countries. The definition varies from country to country, taking into account different quantitative[5] and firm-characteristic[6] variables. Given the scope of this paper we restrict our focus to two important economic zones: the US and the European Union.

The European Union has had a common definition of SMEs since 1996 and this was updated in 2003,[7] probably to take into account the new Basel rules. The number of employees and the annual turnover of a firm are the criteria considered (less than €50 million in sales or less than 250 employees).

In the US there is a special organization (the Small Business Administration) that deals with the politics relating to SMEs and also with their definition based on the North American Industry Classification System. Four criteria are used to identify small-business firms:[8] three generic qualitative rules and one quantitative requirement linked to the industry type. In general, the maximum number of employees is 500 and the average annual receipts should be less than US$28.5 million, but these limits are different for each industry.

[5] The most commonly used are annual turnover, total assets, number of employees, average annual receipts or capital.
[6] Usually no attention is given to legal form but independence from big enterprises, work organization and industry type are often considered.
[7] Commission Recommendation 96/280/EC of April 3, 1996, updated in 2003/361/EC of May 6, 2003, enacted from January 1, 2005.
[8] A small business is one that: is organized for profit; has a place of business in the US; makes a significant contribution to the US economy by paying taxes or using products, materials or labor; and does not exceed the numerical size standard for its industry. For the specific table, see www.sba.gov/size/summary-whatis.html.

Facing all these different criteria, the Basel committee has mainly chosen to follow the annual turnover definition, setting the same general rules to calculate the capital requirements for all the firms (large, medium and small) but ensuring a lighter treatment for small and medium ones (those with an annual turnover of less than €50 million). We believe that this decision is based on the assumption that small firms have a lower default correlation with each other and not because they are considered less risky, in terms of lower expected losses, than the larger firms. Moreover, some SMEs can be classified as retail, but in this case the SME definition does not play any role. The only criterion considered is the bank's exposure (less than €1 million). We conclude that, with this rule, banks are motivated to use automatic decision systems to manage clients with lower exposures regardless of whether they are firms or private individuals, in order to improve their internal efficiency.

2.2 Small and medium-sized enterprise failure

Small and medium-sized enterprise failure rates are very often difficult to track properly. However, in the past few years, considerable research (see, for example, Phillips and Kirchhoff (1989); Watson and Everett (1993); Everett and Watson (1998); Headd (2003)) has been conducted to determine the rates and causation of such failures.

Two of the principle reasons businesses suffer unexpected closures are insufficient capitalization and lack of planning. In the venture capital community it has been found that few, if any, venture capitalists invest their funds into any company that does not have a plan and, in fact, they usually require a business plan to begin the investment process. It is largely because of this that companies in venture-capitalist portfolios have a much higher success rate than those that were financed by banks. Similarly, when investment banks consider a company they promptly look at all the planning documents and financial models applicable to the firm before agreeing to handle the firm as a client. The bank requires three years of taxes, current proof of any income, a financial statement and, if the company is already operating, financials for the company for at least two years. As such, banks take into account only a snapshot of the firm's current economy but do not consider the ability of the applicant to bring the loan to maturity.

When analyzing business failure it is extremely important to distinguish between failure and closure. Watson and Everett (1996) mention that closing firms could have been financially successful but closed for other reasons: the sale of the firm or a personal decision by the owner to accept employment with another firm, to retire or the like. To define failure they create five categories: ceasing to exist (discontinuance for any reason); closing, or a change in ownership; filing for bankruptcy; closing to limit losses; and failing to reach financial goals. Headd (2003) finds that only one-third of new businesses closed under circumstances that owners considered unsuccessful.

We believe that it is essential to carry out this kind of analysis before starting to develop a default prediction model on a sample of small firms. Separating the cases of closure from the ones of failure improves the quality of the available information and of the prediction power of the model, thereby helping to exclude possible outliers from the sample and avoiding biases. In this paper we have taken into account only small-business firms that entered into liquidation, administration or receivership between 2000 and 2007.

2.3 Default prediction methodologies

The literature on default prediction methodologies is substantial. Many authors during the last 40 years have examined several possible realistic alternatives to predict customer default or business failure. The seminal works in this field are by Beaver (1967) and Altman (1968), who develop univariate and multivariate models to predict business failures using a set of financial ratios. Beaver (1967) uses a dichotomous classification test to determine the error rates that a potential creditor would experience if it classified firms on the basis of individual financial ratios as failed or non-failed. He uses a matched sample consisting of 158 firms (79 failed and 79 non-failed) and analyzes 14 financial ratios. Altman (1968) uses an MDA technique to solve the ambiguity problem linked to Beaver's univariate analysis and to assess a more complete financial profile of firms. The analysis draws on a matched sample containing 66 manufacturing firms (33 failed and 33 non-failed) that filed a bankruptcy petition during the period 1946–65. Altman examines 22 potentially helpful financial ratios and selects five that provide, when combined, the best overall prediction of corporate bankruptcy.[9] The variables are classified into five standard ratio categories: liquidity, profitability, leverage, solvency and activity ratios.

For many years thereafter, MDA was the prevalent statistical technique applied to the default prediction models. It was used by many authors (Deakin (1972); Edmister (1972); Blum (1974); Eisenbeis (1977); Taffler and Tisshaw (1977); Altman *et al* (1977, 1995); Micha (1984); Gombola *et al* (1987); Lussier (1995)). However, in most of these studies, authors pointed out that two basic assumptions of MDA are often violated when applied to the default prediction problems.[10] Moreover, in MDA models the standardized coefficients cannot be interpreted like the slopes of a regression

[9] The original z-score model (Altman (1968)) used five ratios: working capital/total assets, retained earnings/total assets, earnings before interest and tax (EBIT)/total assets, market value equity/book value (BV) of total debt and sales/total assets.

[10] Multiple discriminant analysis is based on two restrictive assumptions: that the independent variables included in the model are multivariate normally distributed; and that the group dispersion matrices (or variance–covariance matrices) are equal across the failing and the non-failing group. See Barnes (1982), Karels and Prakash (1987) and McLeay and Omar (2000) for further discussions on this topic.

equation and hence do not indicate the relative importance of the different variables. Considering these problems with MDA, Ohlson (1980) applied for the first time the conditional logit model to the default prediction study.[11] The practical benefits of the logit methodology are that it does not require the restrictive assumptions of MDA and that it allows working with disproportional samples.

From a statistical point of view, logit regression seems to fit well with the characteristics of the default prediction problem, where the dependant variable is binary (default/non-default) and where the groups are discrete, non-overlapping and identifiable. The logit model yields a score between 0 and 1, which conveniently gives the client's probability of default.[12] Lastly, the estimated coefficients can be interpreted separately as the importance or significance of each of the independent variables in the explanation of the estimated probability of default. After Ohlson (1980), most of the academic literature (Zavgren (1983); Gentry *et al* (1985); Keasey and Watson (1987); Aziz *et al* (1988); Platt and Platt (1990); Ooghe *et al* (1995); Mossman *et al* (1998); Charitou and Trigeorgis (2002); Becchetti and Sierra (2002)) used logit models to predict default. Despite the theoretical differences between MDA and logit analysis, studies (see, for example, Lo (1985)) show that empirical results are quite similar in terms of classification accuracy. Indeed, after careful consideration of the nature of the problems and of the purpose of this study we have decided to choose logistic regression as an appropriate statistical technique. Our company-level observations are pooled over time and the covariates are time varying for individual firms until the year of failure. Nam *et al* (2008) show that this logistic estimation is equivalent to a discrete-time duration-dependent hazard model, where the baseline hazard can be related to macroeconomic conditions. We model the baseline hazard using sector-level failure rates.

In order to evaluate the performance of the models we report receiver operating characteristics (ROCs). The ROC curve plots the true positive against the false positive rate as the threshold to discriminate between failed and non-failed firms' changes. The area under the ROC curve (AUC) is a measure of the prediction accuracy of the model, with a value of 1 representing a perfect model. The Gini coefficient and Kolmogorov–Smirnov statistics, commonly used by scoring analysts to evaluate both in-sample and hold-out tests of predictive accuracy, can both be derived from the AUC.[13] We present,

---

[11] Zmijewski (1984) was the pioneer in applying probit analysis to predict default but, until now, logit analysis has given better results in this field.

[12] Critics of the logit technique, including one of the authors of this paper, have pointed out that the specific functional form of a logit regression can lead to bimodal (very low or very high) classification and probabilities of default.

[13] The AUC and the equivalent index, the Gini coefficient, are widely used to measure the performance of classification rules and sidestep the need to specify the costs of different kinds of misclassification. The AUC is a measure of the difference between the score distributions of failed and

in addition, the classification accuracy of our best-fitting models within sample and on the hold-out sample.

3 SMALL-BUSINESS FAILURE PREDICTION MODEL

3.1 Data sample

Our database consists of 5,749,188 sets of accounts for companies that survive in the period 2000–2007 and 66,833 companies that failed in this time period. We retain data from 2006–7 as a test sample (hold-out sample). The accounts analyzed for failed companies are the last set of accounts filed in the year preceding insolvency. For live companies we include accounts for each of the surviving years and estimate a hazard function. In line with other studies we define corporate failure as entry into liquidation, administration or receivership (see Section 2.2). We employ accounting, event, audit and firm-characteristic data to predict the probability of corporate failure of unlisted firms, and we discuss this data in Section 3.2. The breakdown of the sample by data availability is given in part (a) of Table 1 on the next page, which covers financial statements, and part (b) of Table 1 on the next page, which covers firms with limited data. We refer to models using full accounting data as the "SME1 models" and models built using a more limited set of data as the "SME2 models". Table caption still needed.

We have full profit-and-loss account data on over 400,000 companies in each year with between 6,000 and 7,000 failures occurring in each year. The pooled sample gives 2,327,146 non-failed companies and 26,256 failed companies. For companies submitting abridged accounts we have a total of 3,422,042 non-failed companies and 40,577 failed companies.

3.2 Accounting ratios: SME1 model

Altman and Sabato (2007) estimate a model for US SMEs using five financial ratios reflecting dimensions of company profitability, leverage, liquidity, coverage and activity. The final specification, estimated using logistic regression, is reported in Table 2 on the next page.

The UK data set contains the ratios reported by Altman and Sabato and, although we have a wide choice of possible financial characteristics, we are interested to test whether this model can be applied to UK companies using both the US coefficients (Table 2 on the next page) and re-estimations based on the UK sample (see Table 3 on page 12 and Table 4 on page 13).

non-failed companies, while the Gini coefficient is an index that can be calculated as $((2 \times \mathrm{AUC}) - 1)$. The Kolmogorov–Smirnov statistic measures the distance between the two distributions at the optimal cut-off point and is approximately $0.8 \times$ Gini.

TABLE 1 This table needs a caption.

(a) Companies with profit-and-loss data and balance sheet data

Year	Non-failed	Failed	Total	Failed/total
2000	376,015	5,343	381,358	0.01401
2001	374,385	4,835	379,220	0.01275
2002	379,685	4,502	384,187	0.01172
2003	378,094	4,352	382,446	0.01138
2004	384,044	3,902	387,946	0.01006
2005	434,923	3,322	438,245	0.00758

(b) Companies with "abridged" accounts

Year	Non-failed	Failed	Total	Failed/total
2000	433,729	6,377	440,106	0.01449
2001	473,776	6,528	480,304	0.01359
2002	501,670	6,580	508,250	0.01295
2003	571,468	6,678	578,146	0.01155
2004	652,838	7,877	660,715	0.01192
2005	818,357	7,486	825,843	0.00906

Part (a) shows the composition of the development sample used to build the model for SMEs that produce a balance sheet and a profit-and-loss report. Part (b) shows the composition of the development sample used to build the model for SMEs that produce a simplified tax report. In the first column, the year when the financial statement was submitted is shown. The second and third columns show the number of non-failed and failed companies for each financial year, respectively. The fourth column presents the total number of SMEs for each year. The last column shows the annual bad rate.

TABLE 2 Altman and Sabato (2007) US SME model.

Variable	Coefficient
Cash/total assets	0.02
EBITDA/total assets	0.18
EBITDA/interest paid	0.19
Retained earnings/ total assets	0.08
Short-term debt/equity	−0.01
Constant	4.28

In the first column the financial index taken into account is shown. The regression coefficient is presented in the second column. (EBITDA stands for earnings before tax, interest and depreciation.)

3.3 Accounting ratios: SME2 model

Companies that take advantage of reporting exemptions submit "abridged" accounts to the public records. The reporting consists of a modified balance sheet with no profit-and-loss or turnover information. The range of financial ratios available to model insolvency risk is, therefore, quite restricted. We examine the impact of this lack of accounting data on failure prediction and the role played by non-ratio data in predicting bankruptcy. As noted above, the sample of smaller companies contains 3,422,042 non-failed companies and 40,577 failed companies.

Our accounting ratios are selected in the following order. First, since our sample is taken from the UK, we employ variables taken from prior studies into failures of UK companies as set out in Taffler (1984) and Altman and Narayanan (1997).[14] We supplement the variables from UK studies with the variables from the models of Altman (1968) and Zmijewski (1984), and the variables from the model of Ohlson (1980) as analyzed by Begley *et al* (1996) and Hillegeist *et al* (2004).[15] Our variable selection also reflects the importance of working capital for the survival of small firms. The literature on trade credit suggests that smaller firms both extend more credit to customers and take extended credit from suppliers when facing decline and financial stress. Hudson (1986) argues that trade credit forms a large proportion of a firm's liabilities, especially for small firms. He proposes that small-firm bankruptcy is mainly influenced by trade creditors rather than bondholders.[16] Therefore, the trade creditors' decision to force bankruptcy would depend on its customers' cash position (the difference between cash assets and the amount trade creditors are owed), its current indebtedness to the bank, its expected future profits, its liquidation value and interest rates.

There is a large degree of overlap between the financial features of a firm being captured by some of these variables and our modeling process, not least due to multicollinearity considerations, requires us to select between them. Interestingly, many of the working capital cycle variables are not strongly correlated with each other. These considerations lead us to choosing the following accountancy-based variables with which to build our models and eventually to predict which firms will become insolvent and go into bankruptcy procedures:

---

[14] Only one peer-reviewed paper surveyed by Taffler is related to unlisted UK companies, with the remainder relating to listed companies (Taffler (1982)).

[15] We recognize that accounting policies and the institutional environment have changed since many of the studies from which we select our variables were undertaken. We select from a wide range of studies and all of the variables taken represent the distillation of a larger number of variables into those best suited for corporate failure prediction.

[16] Trade creditors would probably act as bondholders in the Bulow and Shoven (1978) model and would have no negotiation or controlling power unless there is a situation where there is only one large trade creditor that has sufficient power to influence decisions.

TABLE 3 SME1 models: z-score and full model.

Variable	Coefficient	Wald	Significance	Coefficient	Wald	Significance
Cash/total assets	−1.487360	2,790.90	0.000000	−1.22627	1,589.06	0.00000
EBITDA/total assets	−0.001980	1,046.34	0.000000	−0.00159	529.15	0.00000
EBITDA/interest paid	−0.002040	62.03	0.000000	−0.00254	102.55	0.00000
Retained earnings/total assets	−0.836940	781.56	0.000000	−0.36787	123.72	0.00000
Short-term debt/equity	0.142100	891.06	0.000000	0.06523	147.65	0.00000
AUDITED	—	—	—	0.56812	1,030.61	0.00000
Audit qualification: severe	—	—	—	0.76862	157.74	0.00000
Audit qualification: going concern	—	—	—	1.03458	982.90	0.00000
Late filing (log of days late)	—	—	—	0.07821	518.77	0.00000
No cash flow statement	—	—	—	0.05697	6.39	0.01148
CCJ number	—	—	—	0.20760	463.03	0.00000
CCJ real value	—	—	—	0.00232	4,520.70	0.00000
Log of age	—	—	—	−0.15921	601.59	0.00000
Age 3–9 years	—	—	—	0.06233	21.77	0.00000
Subsidiary	—	—	—	−0.36864	301.22	0.00000
Subsidiary negative net worth	—	—	—	−0.07076	5.76	0.01641
Size (log)	—	—	—	0.33255	1,056.74	0.00000
Size squared (log)	—	—	—	−0.01122	637.77	0.00000
Industry insolvency	—	—	—	−0.56665	1,628.88	0.00000
Constant	−4.296258	309,527.72	0.000000	−5.83933	5,689.51	0.00000

This table shows models developed for the SMEs that provide balance sheet and profit-and-loss information. The first model includes only the Altman and Sabato (2007) variables and the second also includes the qualitative information. In the first column, the variables entered in the models are presented. The second and fifth columns show the coefficient for each of the variables that entered the model. The third and sixth columns provide the Wald's test value. The fourth and last columns show the significance test value.

TABLE 4 SME2 models: z-score and full model.

Variable	Coefficient	Wald	Significance	Coefficient	Wald	Significance
Retained profit/total assets	−0.093388	1,144.08	0.000000	−0.089649	869.88	0.000000
Quick assets/current assets	−1.091555	3,179.69	0.000000	−0.769366	1,393.34	0.000000
Netcash/networth	−0.051342	216.52	0.000000	−0.042911	157.11	0.000000
Current ratio	−0.095322	990.61	0.000000	−0.047062	316.65	0.000000
Trade creditors/total liabilities	0.208167	150.26	0.000000	0.099292	30.41	0.000000
Trade debtors/total assets	1.569317	7,196.57	0.000000	1.316143	4,498.88	0.000000
Stock/working capital	−0.000046	2.21	0.136838	−0.000073	5.53	0.018708
Change in net worth	−0.001057	768.13	0.000000	−0.000815	469.97	0.000000
Change in RETA	−0.000273	133.71	0.000000	−0.000221	80.76	0.000000
Size (log)	0.303799	11,497.97	0.000000	0.312841	341.16	0.000000
Size squared (log)	—	—	—	−0.001292	2.66	0.102906
Audited	—	—	—	0.144033	88.09	0.000000
Account qualification: severe	—	—	—	0.856334	363.85	0.000000

This table shows models developed for the SMEs that provide limited financial information. The first model includes only financial variables and the second includes also the qualitative information. In the first column, the variables entered in the models are presented. The second and fifth columns show the coefficient for each of the variables that entered the model. The third and sixth columns provide the Wald's test value. The fourth and last columns show the significance test value.

TABLE 4 Continued.

Variable	Coefficient	Wald	Significance	Coefficient	Wald	Significance
Account qualification: going concern	—	—	—	0.493064	50.50	0.000000
Log of age	—	—	—	−0.254680	2,156.62	0.000000
Age 3–9 years	—	—	—	0.024190	5.26	0.021820
Late filing (log of days late)	—	—	—	0.094853	1,252.32	0.000000
Subsidiary	—	—	—	−0.476672	274.06	0.000000
Subsidiary negative net worth	—	—	—	0.165492	15.31	0.000091
Industry insolvency	—	—	—	−0.625116	2,937.54	0.000000
Number of CCJs	—	—	—	0.212898	903.54	0.000000
Real value CCJs	—	—	—	0.001197	7,971.77	0.000000
Constant	−7.554463	36,879.93	0.000000	−6.092687	3,760.99	0.000000
Non-failed: 3,422,042						
Failed: 40,577						

1) capital employed/total liabilities;
2) quick assets/current assets;
3) current assets/current liabilities;
4) total liabilities/quick assets;
5) trade creditors/trade debtors;
6) trade creditors/total liabilities;
7) trade debtors/total assets;
8) inventory/working capital;
9) cash/total assets;
10) net cash/net worth;
11) retained profit/total assets; and
12) short-term debt/net worth.

With respect to leverage variables, a firm's capital employed/total liabilities includes shareholders' funds plus long-term liabilities divided by long-term liabilities and represents the book value of the capital structure of the company. Financially distressed firms would be expected to have larger liabilities relative to shareholders' funds and will therefore have lower values for this variable than healthier entities would.

A number of variables reflect a firm's working capital. The quick assets/current assets variable determines the extent to which current assets consist of liquid assets. The cash/total assets variable expresses cash as a proportion of total assets. The net cash/net worth variable measures net cash as a proportion of net worth. Many firms fail due to a lack of liquid assets and financially distressed firms would therefore be expected to have lower values for these variables. Other variables reflecting the working capital cycle are total liabilities/quick assets, trade debtors/total assets, trade creditors/trade debtors, trade creditors/total liabilities and inventories/working capital. Smaller companies often rely heavily on trade finance from suppliers when bank finance is not available to them. Moreover, small companies extend trade credit to customers as a means of gaining and retaining customers. The use and extension of trade credit makes the business vulnerable to cashflow difficulties.

Retained profit/total assets is a measure of the cumulative profitability of the firm, its leverage and the age of the company. Firms that are unable to accumulate profit from sales will have lower values of this variable. Short-term debt/net worth measures

TABLE 5 Potential non-financial information.

Type and sector	Size and age
Subsidiary	Total assets (size bands)
Owner-managed	Age from incorporation
Family business	Age bands
Leveraged buyout	• “Honeymoon” (<3 years)
Sector risk (failure rate)	• Age risk (3–9 years)
Reporting and compliance	**Operational risk**
Provide full accounts	Court actions to recover debt
Provide cashflow statements	• Debt default values
Audited company	• Number of court actions
Filing history	• “Charges” on assets
• Late accounts	Auditor opinions/qualifications
• Changes in directors	• Severe qualifications
Auditor switching	• Going concern
	• Mild qualifications
	• Scope limitation

the changes in net worth and retained profit/total assets year on year. Financially distressed firms are more likely to have a declining and/or negative net worth. The inclusion of these variables allows us to control for both the level and the direction of net worth and profit.

3.4 The value of non-financial and non-accounting information

A potentially powerful addition to annual financial data available on SMEs is the occurrence of “event” data, such as evidence of a company defaulting on credit agreements and/or trade credit payments or variables representing operational risk, and regulatory compliance, such as whether the firm is late to file its financial statements. Some of these “default events” are available on a monthly basis from a government agency and will enable our model to adjust risk scores more frequently than is possible with annual accounting data. Examples of event data and other potentially predictive information initially explored are listed in Table 5.

A county court judgment (CCJ) arises from a claim made to the court following the non-payment of unsecured debt (usually trade debts). Where the creditor’s claim is upheld by the court, a CCJ is issued. This is an order from the court stating that the debt must be settled. After being issued, either a CCJ is satisfied or it remains outstanding.

The accumulation of CCJs and/or CCJs against companies that are already showing signs of financial distress is likely to be an effective predictor of insolvency. In this study we find that CCJs are better predictors of the likelihood of failure for small companies than for very large companies. This may be due to the fact that certain large companies often "abuse" their bargaining power and are slow payers, forcing creditors to apply to the courts. This is particularly the case when it is relatively easy to dispute invoices arising from trade credit agreements that are "incomplete contracts". County court judgments for firms that have adequate cashflow and liquidity may proxy "operational risk". Nonetheless, individual creditor claims via the court do not represent a bankruptcy risk for very large companies. We also employ the variable capturing the real value of CCJs within the previous 12 months at the end of the accounting year for the last set of accounts.

The second type of compliance information that we employ relates to the timeliness of the filing of accounts. This is represented by the late filing days variable. Unlisted companies have a period of 10 months within which to file accounts following the end of the financial year. The late filing days variable is the number of days following this 10-month period and relates to non-compliance with regulation. A number of reasons, usually quite negative, can cause these delays. The late filing of accounts may be:

1) a deliberate action to delay the publication of unfavorable information in the event that companies face financial difficulties;

2) a by-product of the financial difficulties a firm faces; or

3) a result of the auditors and directors having disagreements regarding a firm's "true" financial position.

In all cases the late filing of accounts is likely to be an indicator of financial distress. We use the log of the number of days late to capture any non-linear effects.

In audited accounts the auditor may document an opinion regarding the financial position of the company. We employ a series of dummy variables to incorporate the data contained in audit reports. The first of these is a dummy variable indicating that the accounts were audited. The variable AUDITED takes a value of 1 where the firm has been audited and 0 otherwise. To qualify for a total audit exemption a company must be a small company with a turnover of less than £5 million and/or assets of less than £2.5 million. Accounts that are not audited therefore belong to smaller firms. The information contained in unaudited accounts is expected to be less reliable than information in audited accounts. Moreover, auditors are likely to be vigilant in identifying likely insolvency and in preventing "technically insolvent" companies from continuing to trade. Another potential variable is evidence of companies switching auditors. This may be indicative of disputes and disagreements with current auditors in relation to the financial health of the company.

For modeling purposes we identify whether the accounts of the company are audited and, if so, whether the auditor has expressed an opinion about the company in the audit report (ie, an audit qualification). We examine the impact of a firm either being unaudited or having a qualified or referred audit report relative to the base case of companies that have no audit qualification. The dummy variables that capture the information contained in audit reports, in descending order of the quality of the report, are as follows: AQREF takes a value of 1 where the audit report is unqualified but referred; AQSCOPE takes a value of 1 where the audit report is qualified due to a scope limitation; AQMILD takes a value of 1 where the audit report is qualified due to mild uncertainties/disagreements; AQGC takes a value of 1 where the audit report has a going concern qualification; AQSEVERE takes a value of 1 where the audit report is qualified due to a severe adverse opinion or disclaimer of opinion.

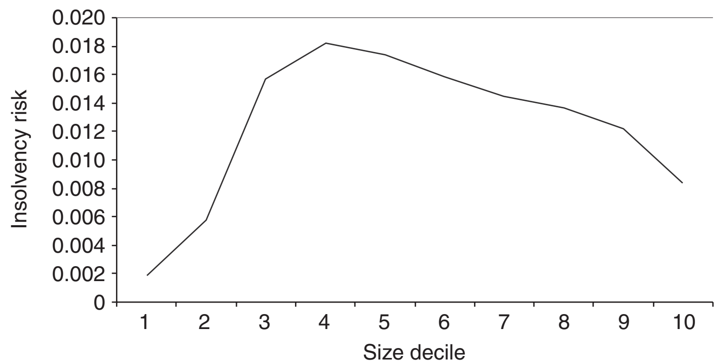
Companies that submit a full set of accounts may sometimes submit a separate cashflow statement along with the profit-and-loss account. We capture this information as a dummy variable "no cash flow statement", which takes the value of 1 if no cashflow statement is provided. We suggest that companies submitting enhanced sets of accounts are likely to be lower risk.

"Liability of newness" theory suggests that a company's risk of exit is highest at the time of start-up and decreases with the age of the company (Stinchcombe (1965)). Hudson (1987), on the other hand, suggests that a newly formed company is most likely to have a "honeymoon period" before being at real risk of failure as it takes time to build up problems and for creditors to get organized into formal insolvency proceedings. Hudson (1987) conducts a survey to understand more about the age, industrial structure and regional structure of liquidated companies using a sample of 1,830 liquidated companies from the period between 1978 and 1981 in the UK. His main finding suggests that young companies form the majority of the liquidated companies and that a company needs at least nine years to be regarded as established (ie, lower the default risk of a start-up). However, he also finds evidence that a newly formed company is most likely to have a "honeymoon period" of around two years.

Following Hudson we employ variables related to the age of the firm as follows: the age of the firm (AGE[17]) at the date of the latest accounts; dummy variables representing firms at particular risk owing to their age, that is, firms between zero and three years of age (AGERISK1 $=$ 1) and firms between three and nine years of age (AGERISK2 $=$ 1). We experiment with combinations of these variables in the model estimation and find that the log of age and AGERISK2 are strongly significant.

[17] The variable AGE is the natural logarithm of the age of the company in years.

FIGURE 1 Relationship between asset size and insolvency risk.



This chart shows the relationship between the insolvency rate and size of company measured by assets. The purpose is to highlight the non-linear relationship between insolvency rate and size in the corporate population. Companies with low asset bases are less likely to be forced into insolvency by creditors since creditors are unlikely to benefit from the process. As the asset base increases insolvency proceedings become more attractive. After a certain threshold point insolvency risk declines with company size.

Our sample includes both non-group companies and subsidiary companies. A subsidiary exists as a separate legal entity and a parent company is protected by limited liability in relation to the liabilities of its subsidiaries. We do not include the parent company, which submits consolidated accounts, since this would lead to the double counting of financial data. Moreover, a subsidiary company may fail as a result of parent-company failure. We are able to identify the parent company of each subsidiary and remove from the failed subsample any subsidiary whose parent company has failed. Following Bunn and Redwood (2003) two dummy variables related to a company being a subsidiary are created. Subsidiary takes a value of 1 where a company is a subsidiary company, and the subsidiary negative net worth variable takes a value of 1 where a company is a subsidiary company and has negative net assets, otherwise subsidiary negative net worth takes a value of 0. A subsidiary company has access to a group's financial and other resources, perhaps leading to a lower likelihood of failure than non-group companies. The group, however, may allow subsidiary companies to fail as part of a wider group strategy. It would potentially be useful to classify other types of ownership including owner-managed and family businesses, and changes in ownership such as buyout activity.

In our models we also control for company size using total asset values. The relationship between asset size and insolvency risk appears to be non-linear, with insolvency risk being an increasing and decreasing function of size (see Figure 1 and Table 6 on the next page). The explanation is that companies with low asset values are unlikely

TABLE 6 Asset values in pounds sterling: size bands applied to all companies.

Size 1	Total assets $<$ 3,000
Size 2	3,000 $<$ Total assets $<$ 50,000
Size 3	50,000 $<$ Total assets $<$ 150,000
Size 4	150,000 $<$ Total assets $<$ 350,000
Size 5	350,000 $<$ Total assets $<$ 700,000
Size 6	700,000 $<$ Total assets $<$ 1,350,000
Size 7	1,350,000 $<$ Total assets $<$ 2,700,000
Size 8	2,700,000 $<$ Total assets $<$ 6,300,000
Size 9	6,300,000 $<$ Total assets $<$ 22,000,000
Size 10	22,000,000 $<$ Total assets

to be pursued through the insolvency process by creditors. This finding is consistent with other studies that find a non-monotone impact of size (see Brüderl *et al* (1992); Falkenstein (2000); Hamerle *et al* (2006)).

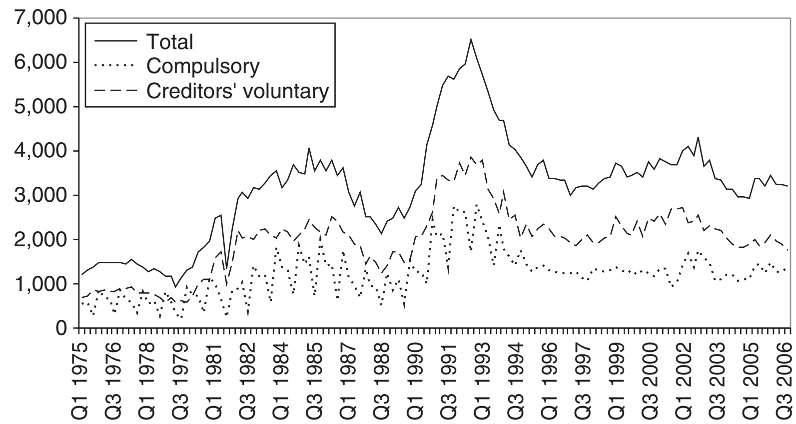
Finally, it is important to control for the macroeconomic conditions facing companies. Although the economic conditions are quite stable over our estimation period there is some sector-level volatility. We are able to control for sector-level risk by calculating the failure rate of the sector in the previous year. Rather than including industrial sector dummy variables we construct a "weight of evidence" variable, which expresses the previous years' sector failure rate as the log odds of failure in each of 51 industrial sectors (INDWOE). This is calculated for each sector using population data on the number of insolvencies in relation to the number of active companies in each sector and acts as a useful proxy for macroeconomic conditions. The INDWOE variable is a useful proxy of the baseline hazard (see, for example, Nam *et al* (2008)).

4 RESULTS

Models are estimated using data pooled from 2000 to 2005, a period of relative stability in the UK economy (see Figure 2 on the facing page). Insolvencies that occurred in 2006 and 2007 are retained for hold-out tests. Financial ratios are corrected for extreme values by restricting the ranges to between the 99th and 1st percentiles.

First, we present the results from the application of the Altman and Sabato (2007) model to the UK sample showing the impact of the addition of non-financial information (SME1 model). In Section 4.2 we estimate a model (SME2) for the companies that file partial accountancy information and rely on, predominantly, balance-sheet ratios to predict insolvency. These companies do not report profit-and-loss account data but have information on retained profit. In Section 5 we present hold-out tests on 2006–7 data.

FIGURE 2 Corporate insolvencies in the UK 1975–2007.



This chart shows the number of insolvencies in the UK corporate sector in the period 1975–2007. The total is broken down by insolvency type, compulsory and voluntary liquidations.

4.1 SME1 model

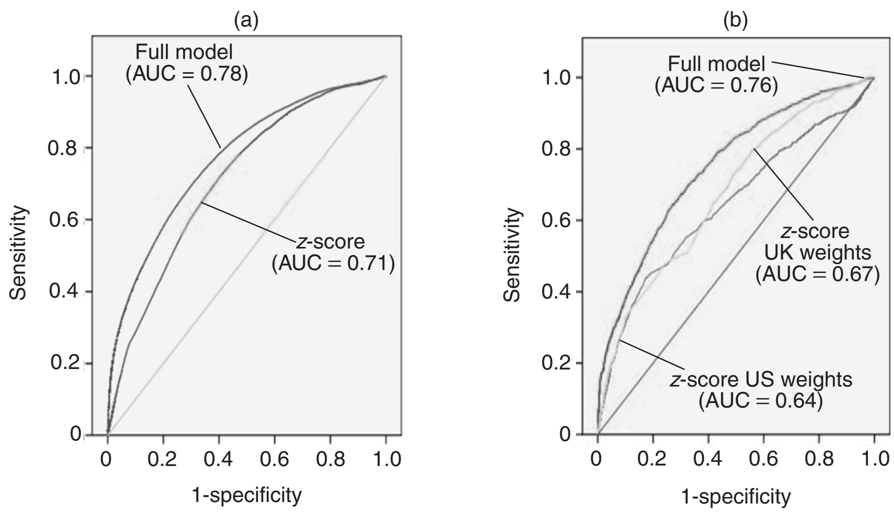
We estimate the model based on the five financial ratios used in the US SME model developed by Altman and Sabato (2007). The model is estimated using logistic regression with 1 being failed and 0 being non-failed, so we expect that a negative coefficient will indicate a reduced risk of insolvency and a positive coefficient will indicate an increased risk of insolvency.[18]

The SME1 model reported in Table 3 on page 12 is built on a sample that includes 2,237,147 non-failed and 26,256 failed companies. The insolvency rate is around 1.2%, which represents the population failure rate for companies that survive for more than one year. The combination of financial ratios specified by Altman and Sabato (2007) is used to model the probability of default. The variables attract the expected signs and are all strongly significant in the equation. Thus, companies with a high ratio of cash to total assets exhibit a lower propensity to failure, as do companies that can adequately cover interest payments on loans out of profits and companies that show higher profit and retained profit to asset ratios. Companies with higher levels of short-term debt to equity are more prone to failure.

The model is re-estimated with the inclusion of a set of non-financial and non-accounting variables. We find that the addition of non-accounting data to the basic

[18] Please note that in this study the dependent variable (defaulted/non-defaulted) is defined in the opposite way from the Altman and Sabato (2007) study. For this reason, the signs of the coefficients are inverted.

FIGURE 3 Receiver operating curves for z-score and full models including qualitative information for the SME1 models.



Receiver operating characteristics curves for (a) within-sample and (b) out-of-sample model performance. Within sample we plot the model performance of the basic z-score model and the curve for the fully enhanced model. The gap between the two curves shows the extent of performance improvement when additional variables are added to the basic z-score. This improvement is also reflected in the AUC statistic. The hold-out sample charts include the basic z-score applied using the US weighting structure as well as the basic z-score with weights re-estimated on the UK sample.

z-score model significantly improves the classification performance. This is shown by the improvements in classification accuracy and associated statistics. The in-sample classification accuracy of the model increases by about 10% (from an AUC of 0.71 to 0.78 including qualitative information) (see Figure 3). The five financial ratios retain their appropriate signs and significance.

We find, as expected, that the age of a company is negatively related to failure propensity, indicating that the longer a company survives the less likely it is to fail. However, our dummy variable representing age 3–9 years is positive and significant. Thus, in line with previous studies, we find that companies in the age bracket of 3–9 years are more vulnerable to failure.

The late filing of accounts is associated with a higher probability of failure. The longer a company takes to file accounts after the year end the more likely it is that the company is encountering difficulties and/or disagreements with the auditors. The no cashflow statement variable is significant and positive, confirming the assertion that companies that submit detailed cashflow statements, thereby volunteering extra information, are generally lower risk. The occurrence of CCJs for the non-payment of trade debt is associated with failure among SMEs, with a decreasing significance the larger the company. We measure CCJ activity for the SMEs in our sample and find that the number and the value of CCJs in the years prior to failure are likely symptoms of financial distress.

In the SME1 model, subsidiaries are less risky than non-subsidiaries. Generally, subsidiaries have access to the financial, and other, resources of the parent company and can survive poor financial performance for longer than non-subsidiaries. Moreover, the parent may have reasons (research and development, tax, etc) for supporting the survival of a subsidiary with recurring negative net worth.

We also find that companies that are audited and that have "audit qualifications" (eg, "severe" or "going concern") are more prone to failure since the auditor is indicating that the long-term viability of the company is in some doubt. The variable AUDITED, which indicates whether or not the company is audited, is positive and significant. This suggests that companies that are subjected to the scrutiny of an auditor are less likely to continue to trade if the company is technically insolvent.

Turning to size we find some interesting results. There is clearly a non-linear relationship between the probability of insolvency and size, as measured by asset values. Descriptive statistics (see Table 7 on the next page and Table 8 on page 25) show an increasing and decreasing relationship between asset values and failure propensity. Clearly, businesses with low asset values are less likely to be pursued through the legal process of insolvency since creditors would have little to gain from the process and these same companies can opt to submit unaudited accounts. We model the size relationship using quadratic terms in the log of total assets. The signs of the coefficients show the expected insolvency/risk-size relationship. The results suggest a threshold level of assets (£350,000) before "legal insolvency" becomes attractive for creditors. Finally, the control for the industry sector is significant and picks up the effects of the average sector-level failure rate on the companies' probability of failure.

4.2 SME2 model

Companies that opt to submit abridged accounts as fulfillment of their reporting requirements are a large and increasing proportion of the limited company population in the UK. For instance, in 2005, of the 1.2 million accounts submitted 765,000 were abridged (60%). The generic models proposed by many researchers to predict insolvency rely on profit and debt ratios that typically cannot be calculated for this large a number of SMEs. In this section we examine the feasibility of building an insolvency-risk model based on the limited information filed in abridged accounts

TABLE 7 Univariate analysis of the SME1 model's variables.

SME1 variables		Mean	Standard deviation
Cash/total assets	Failed	0.1254	0.2291
	Non-failed	0.2649	0.3415
EBITDA/Total assets	Failed	−6.7501	99.1122
	Non-failed	33.0589	131.2077
EBITDA/interest paid	Failed	0.7388	31.4972
	Non-failed	3.6580	33.9341
Retained earnings/total assets	Failed	−0.0421	0.2236
	Non-failed	−0.0021	0.1282
Short-term debt/equity	Failed	0.3399	1.2717
	Non-failed	0.1306	0.7741
Audit qualification: severe	Failed	0.0145	0.1196
	Non-failed	0.0022	0.0469
Audit qualification: going concern	Failed	0.0528	0.2237
	Non-failed	0.0087	0.0929
Late filing (log of days late)	Failed	1.5230	2.0312
	Non-failed	0.8920	1.6183
CCJ number	Failed	0.4101	1.4117
	Non-failed	0.0258	0.2454
CCJ real value	Failed	131.8824	291.8617
	Non-failed	11.3063	89.5378
Age 3–9 years	Failed	0.4326	0.4954
	Non-failed	0.4092	0.4917
Subsidiary	Failed	0.2542	0.4354
	Non-failed	0.2242	0.4171
Log of age	Failed	7.5236	1.0835
	Non-failed	7.5132	1.1303
Subsidiary negative net worth	Failed	0.0873	0.2822
	Non-failed	0.0566	0.2310
AUDITED	Failed	0.5118	0.4999
	Non-failed	0.3373	0.4728
Size (log)	Failed	12.4052	2.8367
	Non-failed	10.7611	4.0716
Size squared (log)	Failed	161.9349	62.7984
	Non-failed	132.3799	77.4382
Industry insolvency	Failed	−0.1254	0.4816
	Non-failed	0.0792	0.4485
No cashflow statement	Failed	0.8529	0.3542
	Non-failed	0.9177	0.2749
	Failed: 24,384		Non-failed: 2,318,764

TABLE 8 Univariate analysis of the SME2 model's variables.

SME2 variables		Mean	Standard deviation
Retained profit/total assets	Failed	−0.3799	1.5458
	Non-failed	−0.1320	1.6297
Quick assets/current assets	Failed	0.7554	0.2993
	Non-failed	0.7816	0.3498
Net cash/net worth	Failed	0.4262	1.7691
	Non-failed	0.7937	2.0760
Current ratio	Failed	1.2853	2.6964
	Non-failed	1.7458	3.3180
Trade creditors/total liabilities	Failed	0.8036	0.2776
	Non-failed	0.7711	0.3601
Trade debtors/total assets	Failed	0.4241	0.3010
	Non-failed	0.2912	0.3156
Stock/working capital	Failed	55.2050	187.6354
	Non-failed	31.7451	131.9738
Size (log)	Failed	11.7055	1.8298
	Non-failed	10.2344	3.2021
Size squared (log)	Failed	140.3662	35.5128
	Non-failed	114.9958	49.7453
Change in net worth	Failed	−18.4766	211.3981
	Non-failed	19.0742	174.0944
Change in RETA	Failed	−56.9872	290.1041
	Non-failed	−3.2406	214.8130
Log of age	Failed	7.3660	1.0260
	Non-failed	7.4480	1.0674
Age 3–9 years	Failed	0.4339	0.4956
	Non-failed	0.4218	0.4938
Late filing (log of days late)	Failed	1.5548	2.0200
	Non-failed	0.8474	1.5711
Subsidiary	Failed	0.0634	0.2437
	Non-failed	0.0586	0.2350
Subsidiary negative net worth	Failed	0.0287	0.1669
	Non-failed	0.0165	0.1272
AUDITED	Failed	0.2170	0.4122
	Non-failed	0.1228	0.3282

TABLE 8 Continued.

SME2 variables		Mean	Standard deviation
Account qualification: going concern	Failed	0.0174	0.1309
	Non-failed	0.0026	0.0507
Account qualification: severe	Failed	0.0073	0.0854
	Non-failed	0.0010	0.0323
Industry insolvency	Failed	−0.2026	0.4547
	Non-failed	0.0132	0.4524
Number of CCJs	Failed	0.5225	1.5078
	Non-failed	0.0304	0.2560
Real value CCJs	Failed	312.2595	609.6046
	Non-failed	22.6622	175.0598

(ie, limited balance sheet information only).[19] After specifying a basic model we then add the range of non-financial and non-accounting data reported above.

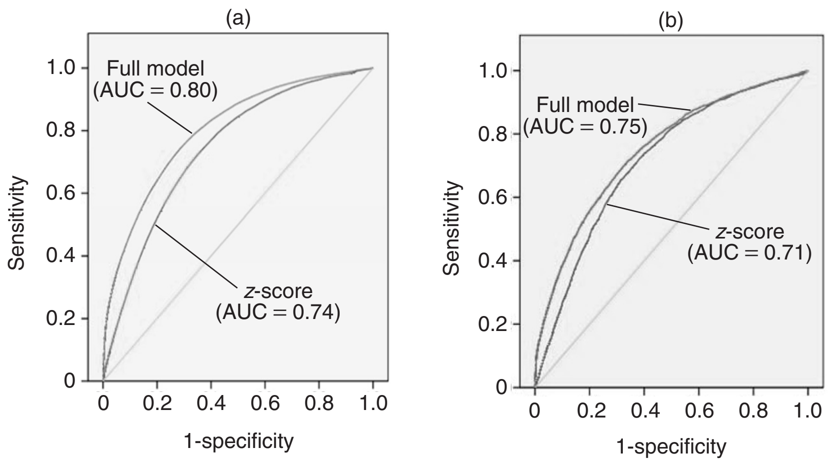
The estimated model is based, again, on a considerable sample size consisting of 3,422,042 non-failed companies and 40,577 insolvent companies. This is reported in Table 4 on page 13. The population failure rate of companies surviving more than one year is around 1.2% during the sample period.

As in the SME1 model, when we add non-financial and compliance information to the basic accounting model the core variables retain their signs and significance and the non-financial variables add value to the model (an AUC of 0.80) with an improvement of over 8% compared with the AUC of the model using only financial information (0.74) (see Figure 4 on the facing page).

The retained profit to total assets ratio is negative and significant, implying that small companies that can accumulate profit from trading are less prone to failure. Having liquidity and cash is associated with a lower probability of failure, measured by various cash ratios. The current ratio can provide conflicting evidence. For many small companies we find that the current ratio actually improves in the financially distressed subsample. However, this effect is almost entirely due to an increase in trade debt relative to short-term borrowing among financially distressed small companies. Related literature on trade credit appears to suggest that financially distressed small companies have higher levels of both trade debt supplied to customers and trade credit obtained from suppliers. The rationale is that small companies may try to boost

[19] Of course, we are able to estimate the abridged model for all companies by excluding the profit-and-loss data from the “full account” company sample but we find that this does not improve the predictive accuracy of the models and we treat full versus abridged as distinct subsamples.

FIGURE 4 Receiver operating curves for z-score and full models including qualitative information for the SME2 models.



The two charts plot ROC curves for (a) within-sample and (b) out-of-sample model performance. Within sample we plot the model performance of the basic z-score model and the curve for the fully enhanced model. The gap between the two curves shows the extent of performance improvement when additional variables are added to the basic z-score. This improvement is also reflected in the AUC statistic. The hold-out sample charts include the basic z-score applied using the US weighting structure as well as the basic z-score with weights re-estimated on the UK sample.

sales by offering credit (emulating their larger competitors) but without having the financial resource to back this strategy. Trade debtors may also increase because customers may avoid paying suppliers that are showing signs of financial difficulty, or it may be that many small companies fail because of late payments by customers (large buyers taking extended credit).

Trade credit as a ratio of total liabilities is higher in the failed subsample than in the non-failed sample. Small companies that are restricted in bank credit may substitute trade credit where possible, taking advantage of the fact that an individual supplier may be unaware of the total amount of trade credit that the company has acquired from other suppliers. As expected, both trade debt to total assets and trade credit to total liabilities are positive and significant. We add three further control variables: log of assets, year-on-year changes in net worth and retained earnings/total assets (RETA).

We again observe the non-linear relationship between asset size and insolvency risk. The size dummies are constructed differently for the SME2 model but the turning point is around the same value as for SME1 (£350,000). Age is negatively associated with failure but, as in the SME1 model, the 3–9 years band attracts a positive and

TABLE 9 Area under ROC comparison of the different models.

	Type of model	
	SME1	**SME2**
Only financial variables		
US weights	0.64	—
UK weights	0.67 (0.71)	0.71 (0.74)
Adding qualitative information		
UK weights	0.76 (0.78)	0.75 (0.80)

This table shows the AUC calculated after plotting the ROC of each one of the three different models on the test sample. The AUC can be interpreted as the average ability of the model to accurately classify defaulters and non-defaulters. The values in the brackets result from the application of the different models on the development sample.

significant sign. Late filing of accounts is associated with a higher probability of failure, as are two audit qualifications, severe qualification (AQSEVERE) and going-concern qualification. As in the previous model, AUDIT is positive and significant. The two variables that measure legal action to recover debts, the number of CCJs and their values, are both positive and strongly significant.

The subsidiary variable is negative and significant but the variable that indicates subsidiaries with negative net worth is positive, suggesting that smaller subsidiaries are not supported by parent companies in the same way that larger subsidiaries are.

5 MODEL VALIDATION

We retained data from 2006 and 2007 in order to undertake hold-out tests for model performance. For these tests we take data from accounts submitted in the first half of 2006 and track all companies that became insolvent in 2007 compared with those that were still alive at the end of 2007. For the SME1 model we identify 236,137 non-defaulted companies and 1,017 that were defaulted by the end of 2007. For the SME2 model we identify 537,865 non-defaulted and 3,040 defaulted companies.

For the SME1 model we generate ROC curves for the out-of-sample test and the classification accuracy of the final models (Table 9). We report the performance of the Altman and Sabato model with the US coefficients applied to the UK data set, the curve for the model re-estimated on UK data and the ROC curve for the model inclusive of the full set of financial and non-financial data. The Altman and Sabato model estimated on US SMEs performs relatively well (AUC of 0.64) when predicting the insolvent companies but its overall performance is affected by its misclassification of non-defaulted companies. When the model is re-estimated on UK data the performance

Sorry that I didn't notice this before Jennifer, but it seems like Table 10 is not mentioned anywhere in the text. Maybe the authors could provide a cross-reference to it? (It was Table 6 in the original submission but there was no cross-reference to Table 6.)

TABLE 10 Classification accuracy rates of the different models.

	Percentage correctly classified		
	Failed (%)	**Non-failed (%)**	**Overall (%)**
SME1 model	76 (76)	73 (75)	74 (76)
SME2 model	77 (80)	73 (76)	75 (78)

This table shows the accuracy rates of the two different models applied to the test sample. The values in parentheses result from the application of the different models to the development sample.

slightly improves (AUC of 0.67). When enhanced with non-financial information the model improves even further, with an AUC of 0.76. Clearly, the addition of these variables significantly improves the overall classification accuracy of the model (13% increase in accuracy on the test sample).

The models developed on abridged accounts show an impressive out-of-sample classification accuracy given the relative lack of information. Clearly, the smaller companies submitting abridged accounts are less complex, or more transparent, than some of their larger counterparts and we note that, although they report less information, insolvency is easier to predict. Again, however, we find an uplift in performance by drawing the ROC curves for models with (0.75) and without (0.71) non-financial information. This is particularly important because in this subsample we are more likely to encounter problems of missing data in the financial statements.

6 CONCLUSIONS

Small and medium-sized enterprises constitute the majority of obligors of banks across Europe and within the US. These companies are not rated either because their financial information is not readily available or because it is provided on an inconsistent basis across companies. Models that can utilize available financial data enhanced with non-financial data could provide a useful tool for underwriters. This is particularly the case as more private company accounts become available.[20]

This study builds upon the previous research of Altman and Sabato (2007), which demonstrated that banks should separate SMEs from large corporates when setting their credit risk systems and strategies. In this paper we confirm the main idea that SMEs require models and procedures that are specifically focused on the SME segment, but we expand our analysis to cover a new geographic area (the UK) using a

[20] Recently in Japan and Scandinavia private company financials have started to become available in XBRL format.

considerable sample, including almost 6 million SMEs. Moreover, for the first time we are able to add non-financial information reflecting company characteristics and aspects of operational risk, such as financial reporting compliance, internal audit and trade credit relationships, as predictive variables of company distress. We improve upon existing models from the literature of SME distress prediction in various ways.

First, we test the Altman and Sabato (2007) SME model on a geographically different sample (UK companies) that includes an extremely high number of small companies (5.8 million) covering a very recent economic period (2000–2007). In doing so we eventually prove the substantial soundness and significant prediction power of this generic SME default prediction model.

Then, for the first time we are able to explore the value added by non-financial information specifically for SMEs. We find that this information, when available, is likely to significantly improve the prediction accuracy of the model by up to 13%. In all models we control for macroeconomic conditions by including a variable that tracks, for each company, the industry sector failure rate for 51 industrial sectors in the previous year. As data relating to company outcomes in the aftermath of the credit crisis and recession become available we will be able to further explore the utility of macroeconomic variables as a proxy for changes in the baseline hazard and the impact on individual firms.

Using the available information we develop a default prediction model for that large cross-section of SMEs for which financial information is very limited (eg, sole traders, professionals, micro-companies, companies that choose simplified accountancy or tax reporting). To the best of the authors' knowledge, in the existing literature solutions to address credit risk management for these clients have never been provided.

Our findings clearly confirm for SMEs what has been found in other studies carried out for large corporations (see, for example, Grunet *et al* (2004)): using non-financial variables as predictors of company failure significantly improves the prediction model's accuracy. However, we believe that this result is even more important for SMEs considering the lack of financial information available for so many of them. Moreover, non-accounting information, such as that used in this study, can be updated frequently, allowing financial institutions to correct their credit decisions in a timely manner. Thus, banks should carefully consider the results of this study when setting up internal systems and procedures to manage credit risk for SMEs.

REFERENCES

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

Altman, E. I. (1993). *Corporate Financial Distress and Bankruptcy*, 2nd edn. Wiley, New York.

Altman, E. I., Marco, G., and Varetto, F. (1994). Corporate distress diagnosis: comparisons using linear discriminant analysis and neural networks (the Italian experience). *Journal of Banking and Finance* **18**, 505–529.

Altman, E. I., and Narayanan, P. (1997). Business failure classification models: an international survey. In *International Accounting*, 2nd edn, Choi, F. (ed). Wiley, New York.

Altman, E. I., and Saunders, A. (1998). Credit risk measurement: developments over the last 20 years. *Journal of Banking and Finance* **21**, 1721–1742.

Altman, E. I., and 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., and Sabato, G. (2007). Modeling credit risk for SMEs: evidence from the US market. *ABACUS* **43**, 332–357.

Aziz, A., Emanuel, D. C., and Lawson, G. H. (1988). Bankruptcy prediction: an investigation of cashflow based models. *Journal of Management Studies* **25**, 419–437.

Barnes, P. (1982). Methodological implications of non-normality distributed financial ratios. *Journal of Business Finance and Accounting* **9**, 51–62.

Basel Committee on Banking Supervision (2004). International convergence of capital measurement and capital standards. Report, Bank for International Settlements.

Beaver, W. (1967). Financial ratios predictors of failure, empirical research in accounting: selected studies 1966. *Journal of Accounting Research* **4**, 71–111.

Becchetti, L., and Sierra J. (2002). Bankruptcy risk and productive efficiency in manufacturing firms. *Journal of Banking and Finance* **27**, 2099–2120.

Begley, J., Ming, J., and Watts, S. (1996). Bankruptcy classification errors in the 1980s: an empirical analysis of Altman's and Ohlson's models. *Review of Accounting Studies* **1**, 267–284.

Blum, M. (1974). Failing company discriminant analysis. *Journal of Accounting Research* **12**, 1–25.

Brüderl, J., Preisendörfer, P., and Ziegler, R. (1992). Survival chances of newly founded business organizations. *American Sociological Review* **57**, 227–242.

Bunn. P., and Redwood, V. (2003). Company accounts based modelling of business failures and the implications for financial stability. Working Paper 210, Bank of England.

Charitou, A., and Trigeorgis, L. (2002). Option-based bankruptcy prediction. Paper presented at *6th Annual Real Options Conference, Paphos, Cyprus, July 4–6, 2002*, pp. 1–25.

Deakin, E. (1972). A discriminant analysis of predictors of business failure. *Journal of Accounting Research* **10**, 167–179.

Eisenbeis, R. (1977). Pitfalls in the application of discriminant analysis in business, finance, and economics. *Journal of Finance* **32**, 875–900.

Everett, J., and Watson, J. (1998). Small business failure and external risk factors. *Small Business Economics* **11**, 371–390.

Falkenstein, E. (2000). RiskCalc$^{\text{TM}}$ for private companies: Moody's default model. Working Paper, Moody's Investors Service, New York.

Fantazzini, D., and Figini, S. (2008). Random survival forest models for SME credit risk measurement. *Methodology and Computing in Applied Probability* **11**, 29–45.

Gentry, J. A., Newbold, P., and Whitford, D. T. (1985). Classifying bankrupt firms with funds flow components. *Journal of Accounting Research* **23**, 146–160.

Gombola, M., Haskins, M., Ketz, J., and Williams, D. (1987). Cash flow in bankruptcy prediction. *Financial Management* **Winter**, 55–65.

Grunert, J., Norden, L., and Weber, M. (2004). The role of non-financial factors in internal credit ratings. *Journal of Banking and Finance* **29**, 509–531.

Hamerle, A., Liebig, T., and Scheule, H. (2006). Forecasting credit event frequency for loan portfolios. *The Journal of Risk* **9**, 75–98.

Headd, B. (2003). Redefining business success: distinguishing between closure and failure. *Small Business Economics* **21**, 51–61.

Hill, P., and Wilson, N., (2007). Predicting the insolvency of unlisted companies. Unpublished manuscript, Credit Management Research Centre, Leeds University.

Hillegeist, S. A., Keating, E. K., Cram, D. P., and Lundstedt, K. G. (2004). Assessing the probability of bankruptcy. *Review of Accounting Studies* **9**, 5–34.

Hudson, J. (1987). The age, regional and industrial structure of company liquidations. *Journal of Business Finance and Accounting* **14**, 199–213.

Karels, G. V., and Prakash, A. J. (1987). Multivariate normality and forecasting of business bankruptcy. *Journal of Business Finance and Accounting* **14**, 573–593.

Keasey, K., and Watson, R. (1987). Non-financial symptoms and the prediction of small company failure: a test of Argenti's hypotheses. *Journal of Business Finance and Accounting* **14**, 335–354.

Litterman, R., and Iben, T. (1991). Corporate bond valuation and the term structure of credit spreads. *Journal of Portfolio Management* **17**, 52–64.

Lo, A. W. (1985). Logit versus discriminant analysis. *Journal of Econometrics* **31**, 151–178.

Lussier, R. N. (1995). A non-financial business success versus failure prediction model for young firms. *Journal of Small Business Management* **33**, 8–20.

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

Micha, B. (1984). Analysis of business failures in France. *Journal of Banking and Finance* **8**, 281–291.

Mossman, Ch. E., Bell, G. G., Swartz, L. M., and Turtle, H. (1998). An empirical comparison of bankruptcy models. *Financial Review* **33**, 35–54.

Nam, C. W., Kim, T. S., Park, N. J., and Lee, H. L. (2008). Bankruptcy prediction using a discrete-time duration model incorporating temporal and macroeconomic dependencies. *Journal of Forecasting* **27**, 493–506.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research* **18**, 109–131.

Ooghe, H., Joos, P., and De Bourdeaudhuij, C. (1995). Financial distress models in Belgium: the results of a decade of empirical research. *International Journal of Accounting* **30**, 245–274.

Peel, M. J., and Peel, D. A. (1989). A multi-logit approach to predicting corporate failure: some evidence for the UK corporate sector. *Omega International Journal of Management Science* **16**, 309–318.

Peel, M. J, and Wilson, N. (1989). Some evidence on discriminating between failing and distressed acquired firms in the UK corporate sector. *Managerial and Decision Economics* **10**, 209–221.

Peel, M. J., Peel, D. A., and Pope, P. F. (1986). Predicting corporate failure: some results for the UK corporate sector. *Omega International Journal of Management Science* **14**, 5–12.

Phillips, B., and Kirchhoff, B. (1989). Formation, growth and survival; small firm dynamics in the US economy. *Small Business Economics* **1**, 65–74.

Platt, H. D., and Platt, M. B. (1990). A note on the use of industry-relative ratios in bankruptcy prediction. *Journal of Banking and Finance* **15**, 1183–1194.

Shanker M. C., and Astrachan, J. H. (2004). Myths and realities: family businesses' contribution to the US economy: a framework for asessing family business statistics. *Family Business Review* **19**, 107–123

Shumway, T. (2001). Forecasting bankruptcy more accurately: a simple hazard model. *Journal of Business* **74**, 101–124.

Stinchcombe, A. L. (1965). Social structure and organizations. In *Handbook of Organizations*, March, J. G. (ed), pp. 142–193. Rand McNally & Company, Chicago, IL.

Taffler, R. J. (1982). Forecasting company failure in the UK using discriminant analysis and financial ratio data. *Journal of the Royal Statistical Society* **145**, 342–358.

Taffler, R. J. (1984). Empirical models for the monitoring of UK corporations. *Journal of Banking and Finance* **8**, 199–227.

Taffler, R. J., and Tisshaw, H. (1977). Going, going, gone: four factors which predict. *Accountancy* **88**, 50–54.

von Stein, J. H., and Ziegler, W. (1984). The prognosis and surveillance of risks from commercial credit borrowers. *Journal of Banking and Finance* **8**, 249–268.

Watson, J., and Everett, J. E. (1993). Defining small business failure. *International Small Business Journal* **11**, 35–48.

Watson, J., and Everett, J. E. (1996). Do small businesses have high failure rates? *Journal of Small Business Management* **34**, 45–62.

Zavgren, C. (1983). The prediction of corporate failure: the state of the art. *Journal of Accounting Literature* **2**, 1–37.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research* **22**, 59–82.