The Value of Qualitative Information in SME Risk Management

Edward I. Altman[a] Gabriele Sabato[b] Nicholas Wilson[c1]

[a] NYU Salomon Center, Leonard N. Stern School of Business, New York University,
44 West 4th Street, New York, NY 10012, USA

[b] Group Credit Risk, Royal Bank of Scotland, Gustav Mahlerlaan 10,
1000EA Amsterdam, The Netherlands

[c] Credit Management Research Centre, Leeds University Business School,
Leeds, LS2 9JT, UK

Abstract

Within the commercial client segment, small business lending is gradually becoming a major target for many banks. The new Basel Capital Accord has helped the financial sector to recognize small and medium sized enterprises (SMEs) as a client, distinct from the large corporate. Some argue that this client base should be treated like retail clients from a risk management point of view in order to lower capital requirements and realize efficiency and profitability gains. In this context, it is increasingly important to develop appropriate risk models for this large and potentially even larger portion of bank assets. So far, none of the few studies that have focused on developing credit risk models specifically for SMEs have included qualitative information as predictors of the company credit worthiness. For the first time, in this study we have available non-financial and 'event' data to supplement the limited accounting data which are 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 qualitative data relating to such variables as 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.

JEL classification: G33, G32, M13
Key words: SME lending; Credit Risk Modeling; Bankruptcy; Small Business failure

[1] Corresponding author. E-mail address: nw@lubs.leeds.ac.uk,
Address: CMRC, Leeds University Business School, LEEDS, UK

1. Introduction

The Basel Capital Accord and the recent financial crisis have provided renewed impetus for lenders to research and develop adequate default/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 definition for major lenders. The literature on the modeling of credit risk for large, listed companies is extensive and gravitates between two approaches (1) the z-score approach of using historical accounting data to predict insolvency (e.g. Altman 1968) and (2) models which 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 (e.g. payment history) and bureau data from the credit reference agencies to develop proprietary models.

In the past, retail lending was mainly synonymous with consumer lending. More recently, following the introduction of Basel II, an increasing number of banks have started to reclassify commercial clients from the corporate area into the retail one. Although this decision may have been originally motivated by expected capital savings (see Altman and Sabato (2005)), financial institutions have soon realized that the major benefits were on the efficiency and profitability side. Banks are also realizing that small and medium sized companies 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, small and medium sized enterprises are the predominant type of business in all OECD economies and typically account for two thirds of all employment. In the UK, unlisted firms make up the majority of firms which 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 as crucial to economic growth and success but, from the lending perspective, research on credit risk management for small companies is relatively scarce. The most likely way to ensure a flow of finance to SMEs is by improving information and developing 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 (see, inter alia, Altman (1968) and (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 work into corporate failure prediction involved determining which accounting ratios best predict failure, employing primarily multiple discriminant analysis or logit/probit models. Neural network models have been also employed to allow for non-linearity (see, inter alia, Altman, Marco and Varetto (1994)), but these models have in turn been criticized for both their lack of a theoretical basis (Altman and Saunders (1998)) and failure to perform better than linear models (Altman et al. (1994)).[2] In most of these accounting ratio based studies, ratios are calculated at a pre-determined time before bankruptcy (usually one year) and as such these models are often referred to as static models. Exceptionally, these studies focus on the use of data other than accounting data, for example von Stein and Ziegler (1984) examine the impact of managerial behaviour on failure. Invariably, this earlier work suffers from only a small sample of failed firms being available for analysis.

Later work on corporate failure prediction involves the application of discrete time hazard models in place of static models.[3] Shumway (2001) argues that the use of static models with multi-period data leads to estimates which are biased, inconsistent and inefficient and he therefore proposes the use of a discrete time hazard model. Further innovations involve the combination of market-based and accounting information (Hillegeist et al. 2004).

[1] The 1993 publication is the second edition of Altman (1983). A third edition, co-authored by Edith Hotchkiss, was published in 2005.
[2] Other models employed include decision tree analysis and Bayesian discriminant analysis.
[3] A further innovation is the use of market based measures in place of accounting ratios (see, inter alia, Hillegeist et al. (2004), Litterman and Iben (1991)). Since our analysis focuses on the prediction of failure across non-listed companies, no market data is available, and Altman and Saunders (1998) argue that major concerns have arisen with the use of a proxy option pricing models for non-listed companies

Recently, Altman and Sabato (2007) apply, 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 compared to large corporate firms, but these organizations should also use scoring and rating systems specifically addressed to the SME portfolio. The lack of any qualitative information about the companies in their sample presents a significant limit forcing them to exclude a relevant portion of small companies without accounting data.

In practice, building credit risk models for private companies are necessarily limited by data availability. Of course, market data are not available for unlisted firms. Further, 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 are not available for SMEs. In recognition of the paucity of data available for many non-listed firms, a paper by Hol (2007) analyses 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 Random Survival Forest (RSF) and compare its performance with a standard logit model.

Recent literature has highlighted the utility of including qualitative variables, 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 qualitative 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). By doing so, we eventually prove the substantial soundness and significant prediction power of our SME default prediction model.

For the first time, we are able to explore the value added of qualitative 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 qualitative information, we develop a default prediction model also for that large part of SMEs for which financial information is very limited (e.g. sole traders, professionals, micro companies, companies that choose simplified accountancy or tax reporting). In the existing literature, solutions to address credit risk management for these clients have never been provided.

The unique database available to us, covers the population of UK companies that have filed accounts during the period 2000 to 2007[4]. The data consists of over 5.8 million records of accounting and other publically 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 small and medium-sized companies based in the UK, have account filing exemptions which mean that the amount of accounting data available for these companies is quite limited. These companies usually represent more than 60% of the total SMEs. Thus, small company[5] accounts include an abbreviated balance sheet and no profit and loss account, and medium-sized[6] company 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.

---

[4] Many previous studies employing UK data are based on samples which contain less than 50 failures. Early studies employing US data similarly contain few failed firms - for example, Altman (1968) employs only 33 failures and Ohlson (1980) 105. More recent US studies are based on larger samples, thus, for example, Hillegeist et al. (2004) employ 756 bankruptcies. As Hillegeist et al. recognise, larger samples will improve the accuracy of the estimation of coefficients and their standard errors. Furthermore, it is difficult to assess industry effects across a small sample. As discussed in Section 3, many prior UK studies are industry specific.

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

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

The remainder of our paper proceeds as follows. In Section 2, we provide an overview of the SME definition, failure definition 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 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 which 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 SME's

Small business lending has been under the attention of researchers and practitioners mainly over the last ten years. Several aspects have been analyzed. From these studies, the definition of small and medium sized enterprises, the definition of SME failure and the modeling techniques that can be applied to predict small business failure are the most relevant for the scope of this study. This Section reviews some of the most important works on these subjects.

2.1 SME definition

We find that there is no common definition of the segment of small and medium sized enterprises across different countries. The definition varies from country to country, taking into account different quantitative[7] and qualitative[8] variables. Considering the scope of this paper, we restrict our focus to two important economic zones (U.S., E.U.).

[7] The most commonly used are: annual turnover, total assets, number of employees, average annual receipts or capital.

[8] Usually no attention is given to the legal form, but independence from big enterprises, work organization and industry type are often considered.

The European Union has had a common definition since 1996 that was updated in 2003[9], probably to take into account also 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 United States there is a special organization (Small Business Administration, or SBA) that deals with the politics relating to SMEs and also with their definition based on the North American Industry Classification System (NAICS). Four criteria are used to identify small business firms[10]: 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 $28.5 million, but these limits are different for each industry.

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 (with annual turnover 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, a part of 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 utilize 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.

[9] Commission Recommendation 96/280/EC of April 3, 1996, updated in 2003/361/EC of May 6, 2003, that replaced the old one from January 1, 2005.

[10] A small business is one that: 1) is organized for profit; 2) has a place of business in the US; 3) makes a significant contribution to the US economy by paying taxes or using products, materials or labor; and 4) does not exceed the numerical size standard for its industry. For the specific table, see http://www.sba.gov/size/summary-whatis.html.

2.2 SME failure

SME failure rates are very often difficult to track properly. However, in the past few years, considerable research (e.g., 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 (VC) community, it has been found that few, if any, VC's invest their funds into any company which does not have a plan and, in point of fact, usually, they require a business plan to begin the investment process. It is largely because of this that companies in VC portfolios have a much higher success rate than those which were financed by banks. Similarly, when investment banks consider a company, they promptly look at all the planning documents and financial models for the firm before agreeing to handle that firm as a client. Rather, 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 created 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. Brian Headd (2003) finds that only one-third of new businesses (33%) closed under circumstances that owners considered unsuccessful.

We believe that this kind of analysis is essential before starting a development of a default prediction model on a sample of small firms. Separating the cases of closure from the ones of failure would improve the quality of the available information and the prediction power of the model 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 voluntary liquidation, administration or receivership between 2000 and 2007.

2.3 Default prediction methodologies

The literature about default prediction methodologies is substantial. Many authors during the last 40 years have examined several possible realistic alternatives to predict customers' default or business failure. The seminal works in this field were Beaver (1967) and Altman (1968), who developed univariate and multivariate models to predict business failures using a set of financial ratios. Beaver (1967) used a dichotomous classification test to determine the error rates a potential creditor would experience if he classified firms on the basis of individual financial ratios as failed or non-failed. He used a matched sample consisting of 158 firms (79 failed and 79 non-failed) and he analyzed 14 financial ratios. Altman (1968) used a multiple discriminant analysis technique (MDA) to solve the ambiguity problem linked to the Beaver's univariate analysis and to assess a more complete financial profile of firms. His analysis drew on a matched sample containing 66 manufacturing firms (33 failed and 33 non-failed) that filed a bankruptcy petition during the period 1946-1965. Altman examined 22 potentially helpful financial ratios and ended up selecting five as providing in combination the best overall prediction of corporate bankruptcy[11]. The variables were classified into five standard ratios categories, including 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), Micha (1984), Gombola et al. (1987), Lussier (1995), Altman et al. (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[12]. Moreover, in MDA models, the standardized

[11] The original Z-score model (Altman, 1968) used five ratios: Working Capital/Total Assets, Retained Earnings/Total Assets, EBIT/Total Assets, Market Value Equity/BV of Total Debt and Sales/Total Assets.

[12] MDA is based on two restrictive assumptions: 1) the independent variables included in the model are multivariate normally distributed; 2) 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 about this topic.

coefficients cannot be interpreted like the slopes of a regression equation and hence do not indicate the relative importance of the different variables. Considering these MDA's problems, Ohlson (1980), for the first time, applied the conditional logit model to the default prediction's study[13]. The practical benefits of the logit methodology are that it does not require the restrictive assumptions of MDA and allows working with disproportional samples.

From a statistical point of view, logit regression seems to fit well the characteristics of the default prediction problem, where the dependant variable is binary (default/non-default) and with the groups being discrete, non-overlapping and identifiable. The logit model yields a score between zero and one which conveniently gives the probability of default of the client[14]. 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 PD. After the work of 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 theoretic 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 the logistic regression as an appropriate statistical technique. In order to evaluate the performance of the models we report receiver operating characteristics (ROC). 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 (K-S), commonly used by scoring analysts to evaluate both in-sample and hold-out tests of predictive accuracy, can both be derived from AUC[15]. We present, in

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

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

[15] The area under the ROC Curve (AUC) and the equivalent index the Gini Coefficient are widely used to measure the performance of classification rules and side-steps the need to specify the costs of different kinds of

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 fail 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 their accounts for each of the surviving years and estimate a hazard function. In line with other studies, we define corporate failure as entry into voluntary 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 these data in the next sections. The breakdown of the sample by data availability is given in Tables 1, Panel A (financial statements) and Panel B (limited data firms). 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'.

Insert Table 1 here

We have full profit and loss account data on over 400,000 companies in each year with between 6 to 7 thousand 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.

---

misclassification. The AUC is a measure of the difference between the score distributions of failed and non-failed companies and the Gini Coefficient is and index which can be calculated as ((2*AUC)-1) and the K-S statistic measures the distance between the two distributions at the optimal cut-off point and is approximately 0.8*Gini.

3.1 Accounting ratios: SME1 model

Altman and Sabato (2007) estimate a model for US SME's using 5 financial ratios reflecting dimensions of company profitability, leverage, liquidity, coverage and activity. The final specification, estimated using logistic regression, is reported in Table 2.

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) and re-estimations based on the UK sample (see Tables 3 and 4).

Insert Table 2 here

3.2 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 and 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.

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 the survey articles of Taffler (1984) and Altman and Narayanan (1997).[16] 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, Ming and Watts (1996) and Hillegeist et al. (2004).[17] 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

[16] Only one peer reviewed article surveyed by Taffler is related to unlisted UK companies, with the remainder being related to listed companies (Taffler (1982)).

[17] 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.

take extended credit from suppliers when facing decline and financial stress. Hudson (1987) 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.[18] 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 the 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 the choice of the following accounting based variables to build our models and eventually to predict which firms will become insolvent and go into bankruptcy procedures.

Capital employed / Total liabilities	Quick assets / Current assets
Current assets / Current liabilities	Total liabilities / Quick assets
Trade Creditors/Trade Debtors	Trade Creditors/Total Liabilities
Trade Debtors/Total Assets	Inventory/working capital
Cash / Total assets	Net cash/Net worth
Retained Profit/Total Assets	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 thus have lower values for this variable than healthier entities.

A number of variables reflect a firm's working capital. *Quick assets/Current assets* determines the extent to which current assets consist of liquid assets. *Cash/Total assets*

[18] Trade creditors would probably act as bondholders of the Bulow and Shoven (1978) model having no negotiation or controlling power unless there is one large trade creditor that can have an influence in decisions.

expresses cash as a proportion of total assets. *Net cash/Net worth* measures net cash as a proportion of net worth. Many firms fail owing to a lack of liquid assets and thus financially distressed firms would 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 cash flow difficulties.

Retained profit/Total assets is a measure of the cumulative profitability of the firm as well as leverage and, finally, 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 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 direction of net worth and profit.

3.3 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 company default on credit agreements and/or trade credit payments or 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 just annual accounting data. Examples of event data and other potentially predictive qualitative information are listed below:

County Court Judgments	Late Filing Days
Audited accounts (y/n)	Audit Report Judgment (e.g. mild, severe, going concern, etc.)
Cash Flow Statement (y/n)	Age of the Firm
Subsidiary (y/n)	Sector

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, a CCJ might be satisfied or remain outstanding. The accumulation of CCJs and/or CCJs against companies that are already showing signs of financial distress are likely to be effective predictors of insolvency. In this study, we find that CCJs are better predictors of the likelihood of failure for small rather than 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. Moreover, 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 County Court Judgments within the previous 12 months at the end of the accounting year for the last set of accounts.

The second type of qualitative information that we employ relates to the timeliness of the filing of accounts. This is represented by the variable, *Late Filing Days*. Unlisted companies have a period of 10 months in which to file accounts following the financial year end. The variable *Late Filing Days* is the number of days following this 10 month period. 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) may be a by-product of the financial difficulties a firm faces, or (3) may be the result of the auditors and directors having disagreements regarding a firm's 'true' financial position. In all cases, it 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. *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 which are not audited therefore belong to smaller firms. The information contained in unaudited accounts would be 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.

For modeling purposes, we identify whether the accounts of the company are audited and if so if the auditor has expressed an opinion about the company in the audit report (i.e. an audit qualification). We examine the impact of a firm being either unaudited or having a qualified or referred audit report relative to the base case of companies that have no audit qualification. The dummy variables which 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 owing to a scope limitation; *AQMILD* takes a value of 1 where the audit report is qualified owing 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 owing to a severe adverse opinion or disclaimer of opinion.

Companies that submit a full set of accounts sometimes may submit a separate cash flow statement along with the P&L account. We capture this information as a dummy variable *No Cash Flow Statement* which takes the value of 1 if no cash flow statement is provided. We suggest that companies submitting enhanced sets of accounts are likely to be lower risk.

Hudson (1987) conducts a survey to understand more about the age, industrial, and regional structure of liquidated companies using a sample of 1,830 liquidated companies between 1978 and 1981 in the UK. His main finding suggests that young companies form the majority of the liquidated companies and a company needs at least around 9 years to be regarded as established (i.e. lower the default risk of a start-up). However, he also points out that a newly formed company is most likely to have a “honeymoon period” of around 2 years before being in real risk. He attributes this finding to the fact that successful firms would distance themselves from the unsuccessful ones over time but even an unsuccessful firm would need some time to build up debts and/or spend the initial investment before finding itself in a financial crisis.

Following Hudson, we employ variables related to the age of the firm as follows: (i) the age of the firm (*AGE*[19]) at the date of the latest accounts, (ii) dummy variables representing firms at particular risk owing to their age, that is, firms between 0 and 3 years of age (*AGERISK1* = 1) and firms between 3 and 9 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.

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 sub-sample 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 *Subsidiary Negative networth* takes a value of 1 where a company is a subsidiary company and has negative net assets, otherwise *Subsidiary Negative Networth* 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.

In our models, we control also 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 Appendix C). The explanation is that companies with low asset values are unlikely to be pursued by creditors through the insolvency process.

Finally, we are able to control for sector level risk by calculating the failure rate in 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

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

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.

4. Results

Models are estimated using data pooled from 2000-2005, a period of relative stability in the UK economy (see Chart 1). 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.

Insert Chart 1 here

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 qualitative information (SME1 model). In paragraph 4.2, we estimate a model (SME2) for the companies that file partial accountant information and rely on, predominantly, balance sheet ratios to predict insolvency. In paragraph 4.3, we present hold-out tests on 2006-2007 data.

4.1 SME1 Model

We estimate the model based on the 5 financial ratios used in the US SME model developed by Altman and Sabato in 2007. The model is estimated using logistic regression with 1=failed; 0=non-failed so that we expect that a negative coefficient indicates a reduced risk of insolvency and a positive coefficient an increased risk of insolvency[20].

The model, reported in Table 3, is built on a sample including 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 more than 1 year. The combination of financial ratios specified by Altman and Sabato (2007) is used to model the probability of default. The

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

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 fail as also companies that can adequately cover interest payments on loans out of profits and show higher profit and retained profit to asset ratios do. Companies with higher levels of short-term debt to equity are more prone to fail.

Insert Table 3 Here

The model is re-estimated with the inclusion of a set of non-financial and non-accounting variables (qualitative information). We find that the addition of non-accounting (qualitative) data to the basic 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 of about 10% (from and AUC of 0.71 to 0.78 including qualitative information) (see Chart 2). The five financial ratios retain their appropriate signs and significance.

Insert Chart 2 Here

We find, as expected, that age of company is negatively related to failure propensity, indicating that the longer a company survives then 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 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 that the company is encountering difficulties and/or disagreements with the auditors. The variable *No Cash Flow Statement* is significant and positive confirming the assertion that companies that submit detailed cash flow statements, and therefore volunteer extra information, are generally lower risk. The occurrence of county court judgments for the non-payment of trade debt is associated with failure amongst SMEs with a decreasing significance the bigger the size of 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 (R&D, tax or other) for supporting the survival of a subsidiary with recurring negative net worth.

We also find that companies that are audited and have 'audit qualifications' (e.g. '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, indicating whether the company is audited or not, is positive and significant. This suggests the 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 Appendix A and B) 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 industry sector is significant and picks up the effects of the average sector level failure rate on the companies probability of failure.

4.2 SME 2 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 amount of SMEs. In this Section, we examine the feasibility of building an insolvency risk model based on the limited information filed in ‘abridged accounts’ (i.e. limited balance sheet information only). 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 and is reported in Table 4. The population failure rate of companies surviving more than 1 year is around 1.2% during the sample period.

Insert Table 4 Here

As in the SME1 model, when we add qualitative information to the basic accounting model the core variables retain their signs and significance and the non-financial variables add value to the model (AUC of 0.80) with an improvement of over 8% compared with the AUC of the model using only financial information (0.74) (see Chart 3).

Insert Chart 3 Here

Retained profit to total assets 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 sub-sample. However, this effect is almost entirely due to an increase in trade debt relative to short-term borrowing amongst 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 sales by offering credit (emulating their larger competitors), but without 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 and/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, the log of assets, and year on year changes in net worth and RETA.

We observe again 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 band 3-9 years attracts a positive and significant sign. Late filing of accounts is associated with a higher probability of failure as are two audit qualifications, severe qualification (AQ_SEVERE) 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 and value of CCJ, 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 in the same way as larger subsidiaries.

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 to those that are still alive as at the end of 2007. For the model SME1, we identify 236,137 non-defaulted companies and 1017 that are defaulted by the end of 2007. For the model SME2, we identify 537,865 non-defaulted and 3040 defaulted companies.

Insert Table 5 Here

For the SME1 model, we generate ROC curves for the out-of-sample test and the classification accuracy of the final models (Table 5). We report the performance of Altman and Sabato's 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) on predicting the insolvent companies, but its overall performance is affected by misclassification of non-defaulted companies. When the model is re-estimated on UK data the performance slightly improves (AUC of 0.67). The model enhanced with qualitative information 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).

Insert Table 6 Here

The models developed on 'abridged accounts' show an impressive out-of-sample classification accuracy given the relative lack of information. Again, we find the uplift in performance by drawing the ROC curves for models with (0.75) and without (0.71) qualitative information.

6. Conclusions

This study builds upon the previous research of Altman and Sabato (2007) that demonstrated that banks should separate small and medium sized firms from large corporates when they are setting their credit risk systems and strategies. In this article, we confirm our main idea that SMEs require models and procedures specifically focused on the SME segment, but we expand our analysis over a new geographic area (UK) using a considerable sample including almost 6 million of SMEs. Moreover, we are able, for the first time, to add qualitative information 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) including an extremely high number of small companies (5.8 million) covering a very recent economic period (2000-2007). By 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 qualitative 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%.

Using available qualitative information, we develop a default prediction model for that large part of SMEs for which financial information is very limited (e.g. 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 for large corporations (e.g. Grunet et al. 2004), that using qualitative 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 that for a large part of them financial information is oftentimes quite limited. Moreover, qualitative information, such as the ones 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 their 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, Vol. 23, pp. 589-609.

Altman, E.I., 1993, 'Corporate Financial Distress and Bankruptcy', 2nd ed. Wiley, New York.

Altman, E.I., G. Marco and F. Varetto, 1994, 'Corporate Distress Diagnosis: Comparisons Using Linear Discriminant Analysis and Neural Networks (The Italian Experience)', *Journal of Banking and Finance,* Vol. 18, pp. 505 - 529.

Altman, E.I., and P. Narayanan, 1997, 'Business Failure Classification Models: An International Survey', in: Choi, F. (Ed.), International Accounting, 2nd ed. Wiley, New York.

Altman, E.I., and A. Saunders, 1998, 'Credit Risk Measurement: Developments Over the Last 20 Years', *Journal of Banking and Finance*, Vol. 21, pp. 1721 – 1742.

Altman, E.I. and G. Sabato (2005) Effects of the new Basel Capital Accord on Bank Capital Requirements for SMEs. *Journal of Financial Services Research,* Vol. 28 (1/3), pp. 15-42.

Altman, E.I. and G. Sabato (2007.) Modeling Credit Risk for SMEs: Evidence from the US Market. *ABACUS* Vol. 43 (3), pp. 332-357.

Aziz, A., Emanuel D.C. and Lawson G.H., 1988, "Bankruptcy prediction – An investigation of cash flow based models". *Journal of Management Studies*, Vol. 25, nr. 5, p. 419-437.

Barnes, P., 1982, "Methodological implications of non-normality distributed financial ratios". *Journal of Business Finance and Accounting*, Vol. 9, nr. 1, Spring 1982, p. 51-62.

Basel Committee on Banking Supervision. 2004. "International Convergence of Capital Measurement and Capital Standards". www.bis.org.

Beaver, W., 1967, "Financial ratios predictors of failure". Empirical Research in Accounting: Selected Studies 1966, *Journal of Accounting Research*, Supplement to Vol. 4, p. 71-111.

Becchetti, L. and Sierra J., 2002, "Bankruptcy risk and productive efficiency in manufacturing firms". *Journal of Banking and Finance*, Vol. 27, p. 2099-2120.

Begley, J., J. Ming and S. Watts, 1996, 'Bankruptcy Classification Errors in the 1980's: An Empirical Analysis of Altman's and Ohlson's Models', *Review of Accounting Studies*, Vol. 1, pp. 267 – 284.

Blum, M., 1974, "Failing company discriminant analysis". *Journal of Accounting Research*, Vol. 12, nr. 1, p. 1-25.

Bunn. P., and V. Redwood, 2003 'Company accounts based modelling of business failures and the implications for financial stability', *Bank of England working paper,* No. 210

Charitou, A. and Trigeorgis L., 2002, "Option-based bankruptcy prediction". Paper presented at 6th Annual Real Options Confernce, Paphos, Cyprus, 4-6 July 2002, p. 1-25.

Deakin, E., 1972, "A discriminant analysis of predictors of business failure". *Journal of Accounting Research*, Vol. 10, nr. 1, Spring 1972, p. 167-179.

Eisenbeis, R., 1977, 'Pitfalls in the Application of Discriminant Analysis in Business', *Journal of Finance*.

Everett, J. and J. Watson, 1998, “Small Business Failure and External Risk Factors”, *Small Business Economics* Vol. 11, pp. 371–390.

Fantazzini, D. and S. Figini (2008), ‘Random Survival Forest Models for SME Credit Risk Measurement’, *Methodology and Computing in Applied Probability*, forthcoming.

Gentry, J.A., Newbold P. and Whitford D.T., 1985, “Classifying bankrupt firms with funds flow components”. *Journal of Accounting Research*, Vol. 23, nr. 1, Spring 1985, p. 146-160.

Gombola, M., Haskins M., Ketz J. and Williams D., 1987, “Cash flow in bankruptcy prediction”. *Financial Management*, Winter 1987, p. 55-65.

Grunet, J., Norden L and Weber M, (2004), ‘The Role of Non-Financial Factors in Internal Credit Ratings’,*Journal of Banking and Finance*, Vol 29 No 2.

Headd, B. 2003, “Redefining Business Success: Distinguishing Between Closure and Failure,” *Small Business Economics*, v. 21, pp. 51-61.

Hillegeist, S.A., E.K. Keating, D.P. Cram, and K.G. Lundstedt, 2004, ‘Assessing the Probability of Bankruptcy’, *Review of Accounting Studies*, Vol. 9, pp. 5 – 34.

Hudson , J 1987, “The Age, Regional and Industrial Structure of Company Liquidations”, *Journal of Business Finance and Accounting,* Vol. 14, nr. 2, pp. 199-213.

Karels, G.V. and Prakash A.J. 1987. “Multivariate normality and forecasting of business bankruptcy”. *Journal of Business Finance & Accounting*, Vol. 14, nr. 4, Winter 1987, p. 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*, Vol. 14 (3).

Litterman, R. and T. Iben, 1991, ‘Corporate Bond Valuation and the Term Structure of Credit Spreads’, *Journal-of-Portfolio-Management* Vol. 17, pp. 52-64

Lo, Andrew W., 1985, “Logit versus discriminant analysis”. *Journal of Econometrics*, Vol. 31, p. 151-178.

Lussier, R.N., 1995, “A non-financial business success versus failure prediction model for young firms”. *Journal of Small Business Management*, Vol. 33, nr. 1, p. 8-20.

Merton, R.C., 1974. ‘On the Pricing of Corporate Debt: The Risk Structure of Interest Rates’, *Journal of Finance*, Vol. 29, pp. 449-470.

Micha, B., 1984, “Analysis of business failures in France”. *Journal of Banking and Finance*, Vol. 8, p. 281-291.

Mossman, Ch.E., Bell G.G., Swartz L.M., and Turtle H., 1998, “An empirical comparison of bankruptcy models”. *The Financial Review*, Vol. 33, nr. 2, p. 35-54.

Ohlson, J.A., 1980, ‘Financial Ratios and the Probabilistic Prediction of Bankruptcy’ *Journal of Accounting Research*, Vol. 18, pp. 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*, Vol. 30, p. 245-274.

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(1). pp5-12.

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(4). pp.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*, Vol. 10 No. 3, pp. 209-221.

Phillips, B. and B. Kirchhoff 1989, "Formation, Growth and Survival; Small Firm Dynamics in the U.S. Economy," *Small Business Economics*, v.1, n.1, pp. 65-74.

Platt, H.D. and M.B. Platt 1990, 'A Note on the Use of Industry-relative Ratios in Bankruptcy Prediction', *Journal of Banking and Finance*, Vol. 15, pp. 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* Vol 19, Issue 2 p107-123

Shumway, T., 2001, 'Forecasting Bankruptcy More Accurately: A Simple Hazard Model', *Journal of Business*, Vol. 74, pp. 101 - 124

Taffler, R.J., 1982, "Forecasting company failure in the UK using discriminant analysis and financial ratio data". *Journal of the Royal Statistical Society*, Vol. 145, Part 3, p. 342-358.

Taffler, R.J., 1984 'Empirical Models for the Monitoring of UK Corporations', *Journal of Banking and Finance*, Vol. 8, pp. 199 – 227.

Taffler, R.J. and Tisshaw H., 1977, "Going, Going, Gone - Four Factors Which Predict". *Accountancy*, Vol. 88, March 1977, p. 50-54.

von Stein, J.H. and W. Ziegler, 1984, 'The Prognosis and Surveillance of Risks from Commercial Credit Borrowers', *Journal of Banking and Finance*, Vol. 8, pp. 249 – 268.

Watson, J. and J. Everett 1993, "Defining Small Business Failure", *International Small Business Journal*, Vol. 11, No. 3, pp. 35-48.

Watson, J and Everett, J.E. 1996, "Do small businesses have high failure rates?", *Journal of Small Business Management*, Vol 34, No. 4, pp.45-62.

Zavgren, C, 1983, "The prediction of corporate failure: the state of the art". *Journal of Accounting Literature*, Vol. 2, p. 1-37.

Zmijewski, M. E.,1984, 'Methodological Issues Related to the Estimation of Financial Distress Prediction Models', *Journal of Accounting Research,* Vol. 22, pp. 59–82.

***Table 1.* Panel A: Companies with Profit & Loss and Balance Sheet Data**

This table 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. In the first column, the year when the financial statement was submitted is shown. The second and third column show the number of non-failed and failed companies for each financial year respectively. The fourth column presents the total number of SME for each year. The last column shows the annual bad rate.

Year	Non-failed	Failed	Total	Failed/Total
2000	376015	5343	381358	0.01401
2001	374385	4835	379220	0.01275
2002	379685	4502	384187	0.01172
2003	378094	4352	382446	0.01138
2004	384044	3902	387946	0.01006
2005	434923	3322	438245	0.00758

***Table 1.* Panel B: Companies with 'Abridged' Accounts**

This table 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 column show the number of non-failed and failed companies for each financial year respectively. The fourth column presents the total number of SME for each year. The last column shows the annual bad rate.

Year	Non-failed	Failed	Total	Failed/Total
2000	433729	6377	440106	0.01449
2001	473776	6528	480304	0.01359
2002	501670	6580	508250	0.01295
2003	571468	6678	578146	0.01155
2004	652838	7877	660715	0.01192
2005	818357	7486	825843	0.00906

Table 2. **Altman and Sabato (2007) US SME Model**
This table shows the Altman and Sabato (2007) model developed for US SMEs. In the first column, the financial index taken into account is shown. In the second column, the regression coefficient is presented.

Variable	Coefficient
Cash/Total Assets	0.020000
EBITDA/Total Assets	0.180000
EBITDA/Interest paid	0.190000
Retained Earnings/Total Assets	0.080000
Short Term Debt/Equity	-0.010000
Constant	4.280000

***Table 3.* SME1 Models: Z-score and full model**

This table shows models developed for the SMEs that provide balance sheet and P&L information. The first model includes only the Altman and Sabato (2007) 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 column show the coefficient for each of the variables that entered the model. The third and sixth column provide the Wald's test value. The fourth and last column show the significance test value.

Variable	Coefficient	Wald	Sig.	Coefficient	Wald	Sig.
Cash/Total Assets	-1.487360	2790.90	0.000000	-1.22627	1589.06	0.00000
EBITDA/Total Assets	-0.001980	1046.34	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	1030.61	0.00000
Audit Qualification- Severe				0.76862	157.74	0.00000
Audit Qulaification - 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	4520.70	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 networth				-0.07076	5.76	0.01641
Size (log)				0.33255	1056.74	0.00000
Size squared (log)				-0.01122	637.77	0.00000
Industry Insolvency				-0.56665	1628.88	0.00000
Constant	-4.296258	309527.72	0.000000	-5.83933	5689.51	0.00000
Non-Failed = 2318764						
Failed = 24384						

Table 4. **SME2 Models: Z-score and full model**

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 column show the coefficient for each of the variables that entered the model. The third and sixth column provide the Wald's test value. The fourth and last column show the significance test value.

Variable	Coefficient	Wald	Sig.	Coefficient	Wald	Sig.
Retained Profit/Total Assets	-0.093388	1144.08	0.000000	-0.089649	869.88	0.000000
Quick Assets/Current Assets	-1.091555	3179.69	0.000000	-0.769366	1393.34	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	7196.57	0.000000	1.316143	4498.88	0.000000
Stock/Working Capital	-0.000046	2.21	0.136838	-0.000073	5.53	0.018708
Change in NetWorth	-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	11497.97	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
Account Qualification _Going Concern				0.493064	50.50	0.000000
Log of Age				-0.254680	2156.62	0.000000
Age 3- 9 years				0.024190	5.26	0.021820
Late Filing (log of days late)				0.094853	1252.32	0.000000
Subsidiary				-0.476672	274.06	0.000000
Subsidiary Negative networth				0.165492	15.31	0.000091
Industry insolvency				-0.625116	2937.54	0.000000
Number CCJs				0.212898	903.54	0.000000
Real value CCJs				0.001197	7971.77	0.000000
Constant	-7.554463	36879.93	0.000000	-6.092687	3760.99	0.000000
Non Failed = 3,422,042 Failed = 40,577						

***Table 5.* AUC comparison of the different models**

This table shows the Area Under the Curve (AUC) calculated after plotting the Receiver Operating Characteristic (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.

		Type of model	
		SME1	SME2
Only financial variables	US weights	0.64	n.a.
	UK weights	0.67 (0.71)	0.71 (0.74)
Adding Qualitative info	UK weights	0.76 (0.78)	0.75 (0.80)

***Table 6.* Classification accuracy rates of the different models**
This table shows accuracy rates of the two different models applied to the test sample. The values in the brackets result from the application of the different models on the development sample.

	Percentage Correctly Classified		
	Failed	Non-failed	Overall
SME 1 Model	76% (76%)	73% (75%)	74% (76%)
SME 2 Model	77% (80%)	73% (76%)	75% (78%)

***Chart 1.* 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.

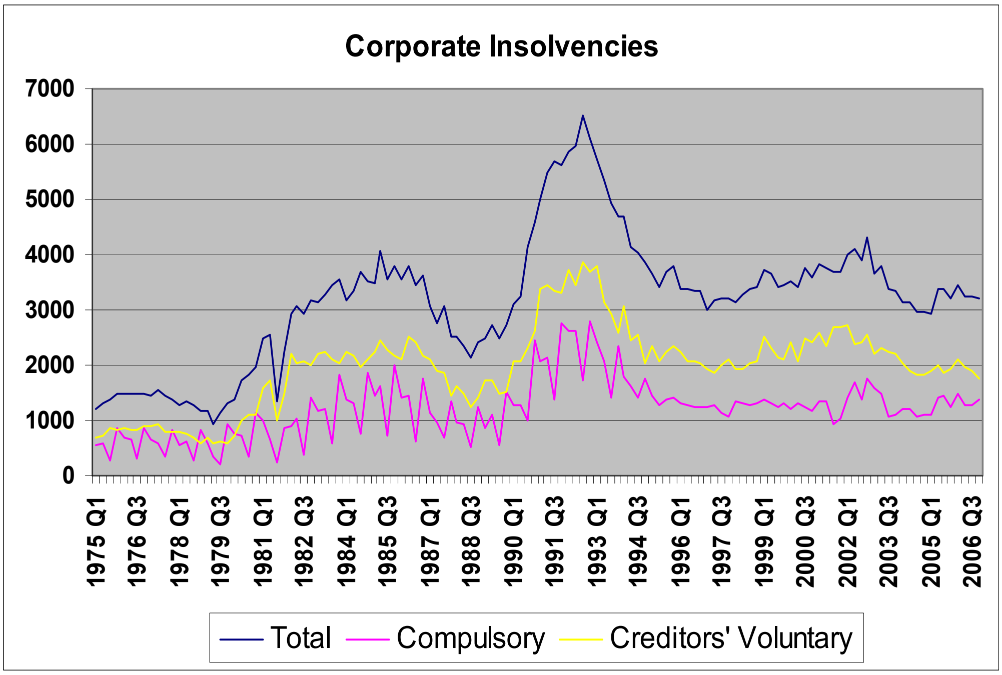


Chart 2. **Receiver Operating Curves for Z-score and full models including qualitative information**

The two charts plot ROC curves for within- and 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 (Area Under the Curve) 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.

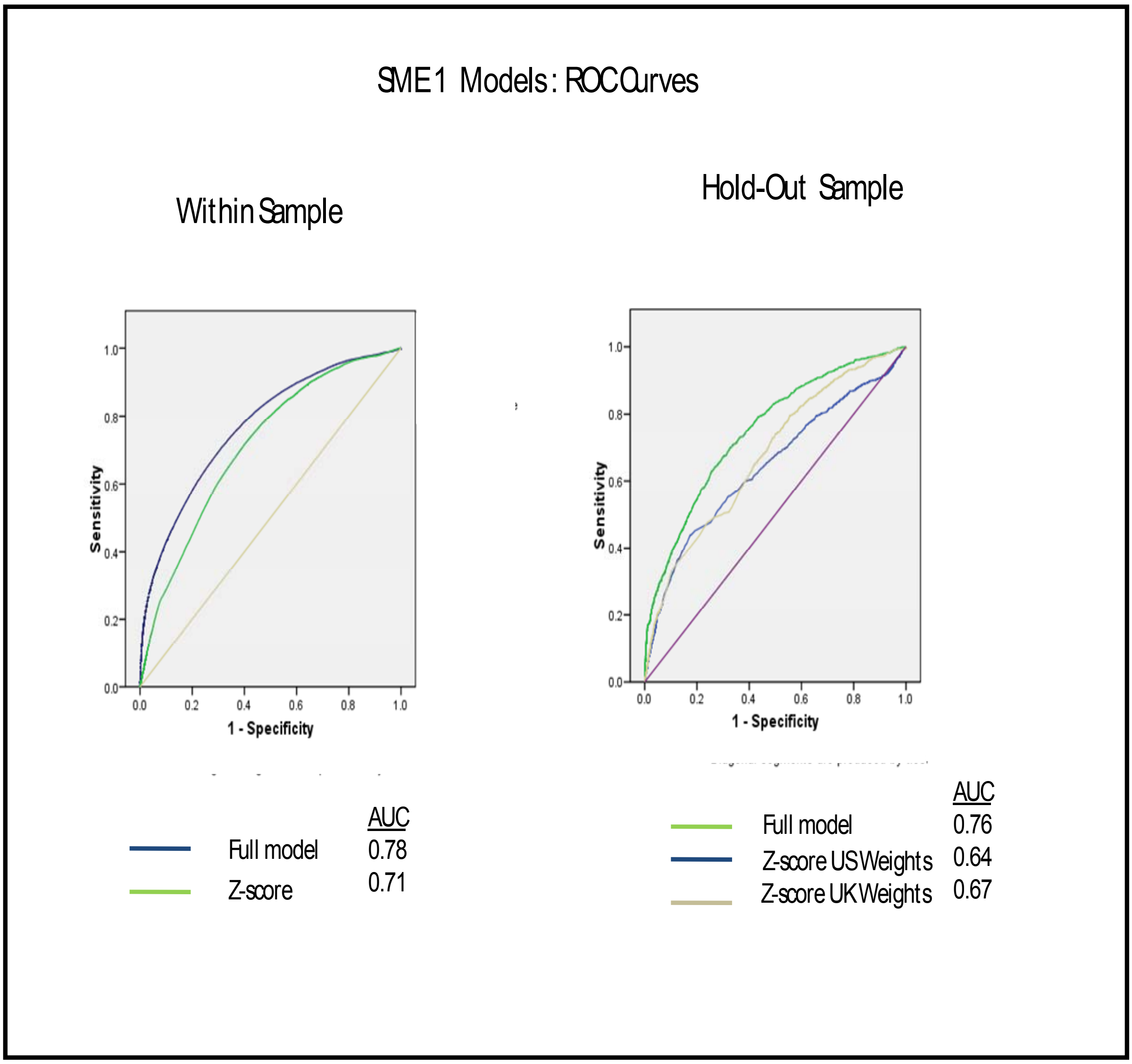
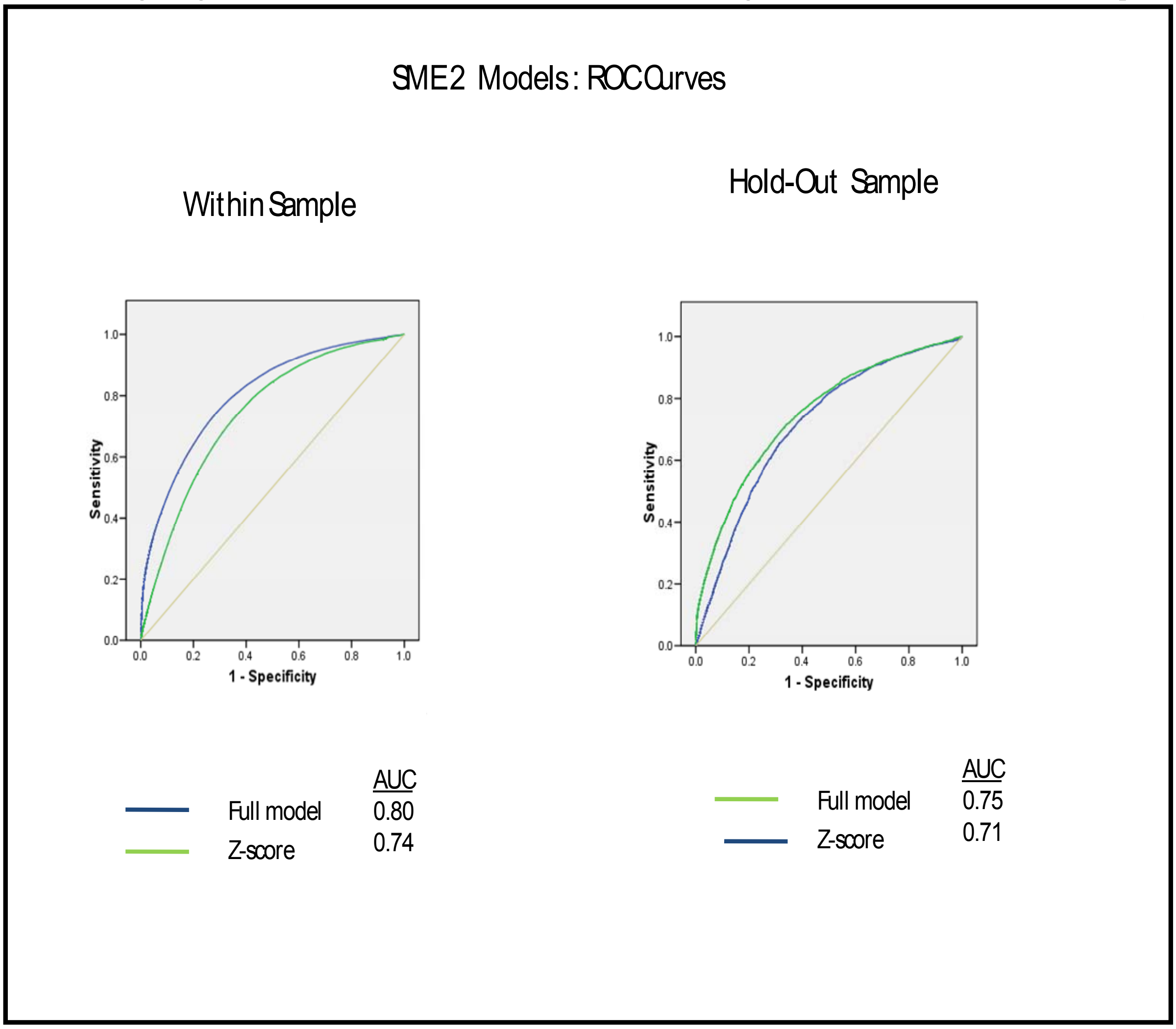


***Chart 3.* Receiver Operating Curves for Z-score and full models including qualitative information**

The two charts plot ROC curves for within- and 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 (Area Under the Curve) 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.



Appendix A: Univariate analysis of the SME1 model's variables

SME 1 Variables

		Mean	Std. 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 Qulaification - 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 networth	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 Cash Flow Statement	Failed	0.8529	0.3542
	Non Failed	0.9177	0.2749

Failed = 24384 Non Failed= 2318764

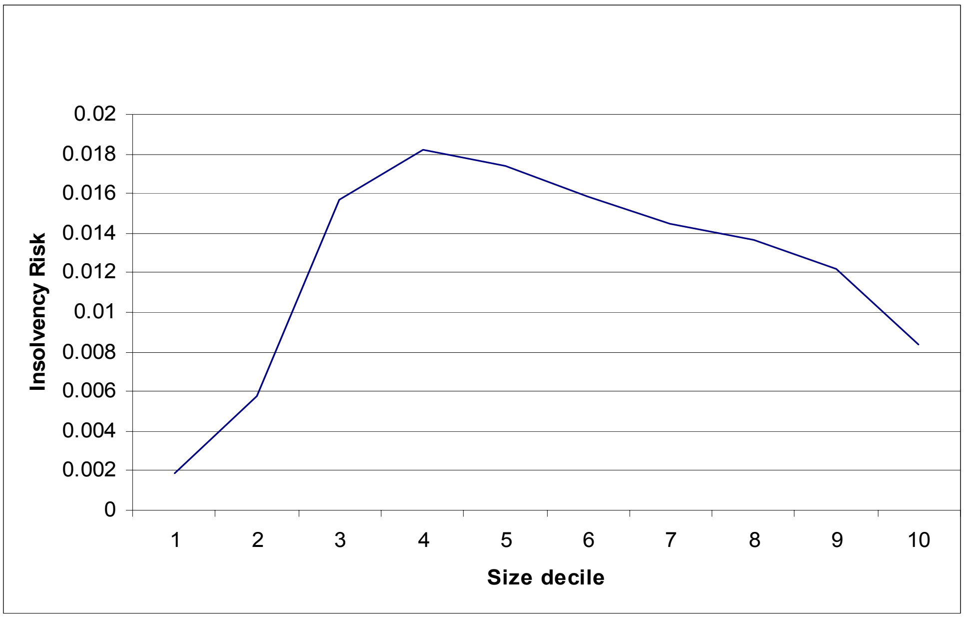
Appendix B: Univariate analysis of the SME2 model's variables

SME 2 Variables

	Status	Mean	Std. 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
NetCash/Networth	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 NetWorth	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 networth	Failed	0.0287	0.1669
	Non Failed	0.0165	0.1272
Audited	Failed	0.2170	0.4122
	Non Failed	0.1228	0.3282
Account Qualification _Going C	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 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

Appendix C: 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.



Asset Values in £'s – size bands applied to all companies

SIZE1	=	Total Assets < 3,000
SIZE2	=	3,000 < Total Assets < 50,000
SIZE3	=	50,000 < Total Assets < 150,000
SIZE4	=	150,000 < Total Assets < 350,000
SIZE5	=	350,000 < Total Assets < 700,000
SIZE6	=	700,000 < Total Assets < 1,350,000
SIZE7	=	1,350,000 < Total Assets < 2,700,000
SIZE8	=	2,700,000 < Total Assets < 6,300,000
SIZE9	=	6,300,000 < Total Assets < 22,000,000
SIZE10	=	22,000,000 < Total Assets