Predicting Unlisted SMEs' Default: Incorporating Market Information on Accounting-based Models for Improved Accuracy

Andrikopoulos, P & Khorasgani, A

Predicting Unlisted SMEs' Default: Incorporating Market Information on Accounting-based Models for Improved Accuracy

Panagiotis Andrikopoulos[1,#] and Amir Khorasgani[1]

Abstract

The risk associated with lending to small businesses has become more important since regulations started obliging banks to use separate procedures in assessing SMEs' credit worthiness. However, current accounting-based models for SMEs do not account for the impact of market information on default prediction. We fill this gap in the literature by introducing a hybrid default prediction model for unlisted SMEs that uses market information of listed SMEs (comparable approach) alongside existing accounting information of unlisted SMEs. Our results suggest that the accuracy of this default prediction modelling approach in the hold-out sample, during the period of the financial crisis 2007-09 and for the entire sample-period, improves considerably. We conclude that the proposed hybrid model is a good replacement for existing standard accounting-based methods on SMEs' default prediction.

[1]*Coventry Business School, Coventry University, Priory Street, Coventry, CV1 5FB, UK*
[#]*Corresponding Author Email: p.andrikopoulos@coventry.ac.uk*

1. Introduction

Credit risk refers to the risk created by unexpected changes in the credit quality of a counterparty or issuer and its quantification is one of the major challenges in modern finance.[1] Traditionally, it is measured as the firm's likelihood of default on its contractual or required obligations and the monetary losses imposed if such default occurs. Although SMEs are the most active economic units representing the backbone of a nation's economy, due to their special characteristics, their credit and operational risks are perceived to be higher. These risks are especially prevalent during periods of prolonged financial crises as they pose a significant threat to the real economy given the potential negative impact on companies' profits, sales and investment (Claessens et al., 2012).

In a recent study, Gupta and Gregoriou (2015) show that the proportion of US SMEs under financial distress in their sample had increased from 19% in 1990 to 31.61% in 2013. At the same time, the proportion of actual bankruptcies had fallen from 1.39% to 0.5% for the corresponding period with micro-SMEs being the companies worst affected.[2] A similar picture is portrayed in the UK where according to the Office for National Statistics (ONS), the average death rate for new businesses in the period 2009-14 is reported as 10.35% with a five-year survival rate of only 41.7% (ONS, 2015). In an attempt to explain this trend, Guariglia et al. (2016) demonstrate a statistically significant link between debt-financing cost and corporate survival rates for young and non-exporting firms, especially during the period of the 2007-09 financial crisis when borrowing rates had increased dramatically. This reduced survival rate for SMEs was amplified by the considerable difficulty in obtaining the necessary finance for their operations. Government statistics in the immediate aftermath of

---

[1] Since it is not certain whether the firm's obligation will be fulfilled or not, the potential borrower's creditworthiness affects all aspects of the firm's cost of capital, the lending decision, the credit spread and the prices and hedge ratios of relevant credit derivatives.

[2] Of all the bankruptcies reported in this study, 40% were attributed to the micro-SMEs sample. A similar pattern is reported for the case of financially distressed firms where micro-SMEs comprise 41.49% of the total sample.

the 2007-09 financial crisis show that only 74% of those SMEs seeking finance in the year 2010 indeed managed to obtain some form of it as compared to a rate of 90% in the period 2007/08 (BIS Economics Paper, 2012). This decreased ability of SMEs to access short-term capital is mainly attributed to the supply of bank lending as banking and financial institutions became more risk averse as well as being legally obliged, by new financial services regulations, to increase their liquidity by holding more capital.[3] Extant literature shows that unlike their large-size counterparts, SMEs' applications for financing tend to be rejected more frequently, with the evidence for this being consistent across the entire spectrum of SME activity, industrial classification and their ability to innovate (Lee et al, 2015).

The severity of this problem in the UK setting has been extensively highlighted over the years in a series of governmental reports. For example, Cruickshank (2000) suggests the existence of systemic problems affecting the quality of lending services for SMEs such as the possible overcharging for such services by banks, the lack of available information regarding alternative banking products, and the existence of significant weaknesses in the systems of redress when things go wrong for the borrowing firms. As the report concludes, these problems are exacerbated by the presence of a complex monopolistic structure in the UK banking environment (Cruickshank, 2000, pp.161-167). A more recent joint-study by the Competition and Markets Authority (CMA) and the Financial Conduct Authority (FCA) further supports such claims. It points out that high barriers to entry and expansion in the supply of business current accounts (BCAs) and general business loans to SMEs for newer and smaller credit providers are a major weakness of the UK bank lending system (CMA & FCA, 2014). This study also reports a significant widening of the difference in the interest rate charged by banks to SMEs for term loans and the Bank of England (BoE) base rate. For example, from an annual average of less than 2.5% in the period prior to the financial crisis,

---

[3] This is an inevitable consequence of the credit crunch period and the global economic crisis that followed.

interest rates have increased to 4% or more in the period 2008-2012. Hence, all prior policy literature unanimously highlights the need for improving market competition in UK banking. Allowing smaller credit providers to enter the market will most certainly improve lending services, thereby reducing both the cost of such services for SMEs and the rate of insolvencies for such firms.

This large increase in SME insolvencies during and after the 2007-09 financial crisis, alongside the introduction of the Basel II capital requirements for banks, have led to a renewed interest in the SME credit risk assessment literature with the main focus on accounting-based default prediction models designed specifically for SMEs.[4] Nonetheless, although these models are easy to use in a practical sense, they ignore the important role of market factors in predicting potential default events. In terms of credit risk quantification using market-based information, one of the most widely-used models in academic literature and practice alike is the one introduced by Merton (1974) and further developed by the KMV Corporation.[5] Although this model has significant advantages over the accounting-based approaches, including better predictability during financial crises and over short-term horizons, the unavailability of market information in the case of unlisted companies deems it inapplicable for the majority of SMEs (Richardson et al., 1998; Bilderbeek and Pompe, 2005; Lin et al., 2007). To the best of our knowledge, there is no prior literature regarding the effect of market-based factors on the accuracy of the unlisted SMEs' default prediction models and especially no elaborate default prediction models that can use market information for predicting their potential default.

Hence, our study addresses this important gap in the literature by introducing a new modelling approach that can utilise market information for predicting unlisted SMEs' default

---

[4] According to Dullmann and Koziol (2013), since small firms are more likely to default due to their idiosyncratic risk, accounting information is the essential tool in SME default prediction models

[5] This study follows the modelling approach of Chen et al. (2010). In line with their study we also refer to this modelling approach as the Merton-KMV model.

using an average sample of 181 UK listed SMEs (L-SMEs thereafter) and 19,681 unlisted SMEs (U-SMEs thereafter) over the period 2004-2013. This is accomplished by combining market and accounting information in a way that takes into account the association between U-SMEs' accounting ratios and the Merton's distance-to-default (DD) for L-SMEs. Our hybrid default prediction model exhibits superior predictive power compared to its existing accounting-based counterparts when tested across our entire sample of U-SMEs.

The rationale behind this approach is simple. Prior studies in the field of corporate finance highlight the benefits of adopting a comparability method in equity valuation of unlisted/private companies using industry-level data (Alford, 1992; McCarthy, 1999; Baker and Ruback, 1999).[6] In a similar manner, we show that such a method has important benefits in default prediction. If the use of a market-based valuation approach is a reliable way for deriving firm value for unlisted/private companies (Alford, 1992; Baker and Ruback, 1999), then there is no theoretical reason why this approach cannot also be used for the purpose of default prediction in the case of such firms. Bhojraj and Lee (2002) further suggest that the use of market-based valuation multiples usually functions as a "satisficing" device for the professional community, trading off methodological complexity and completeness for the purpose of convenience and cost efficiency. Likewise, from the perspective of a large and highly-diverse bank engaging primarily in transaction lending technologies, lending to informationally 'opaque' SMEs poses significant risks and costs. Assuming a decision to lend to a L-SME or an U-SME is mutually exclusive, a bank will most likely be inclined to lend to the former type of company as the cost of accessing all relevant 'soft' information for the U-SMEs can be relatively high.[7] Our proposed approach solves this problem, by allowing U-SMEs to be treated as if they are of a 'quasi-listed' status and be judged in terms of credit ability on an equal basis to the former type of firms. Furthermore, under conditions of market

---

[6] This approach is typically used for the purposes of fundamental analysis or multiple-based valuation in IPOs.
[7] Such higher costs typically lead to higher interest rates on borrowing for the case of SMEs (Baas and Schrooten, 2006).

efficiency[8] the use of market-based information can be considered as a credible source of anticipated economic conditions, as the latter will affect not only the L-SMEs but the U-SMEs as well.[9]

Our results indicate that the new methodological approach can forecast U-SMEs' default better than the traditional way of empirically using a set of accounting ratios, as the use of the proxied market information significantly increases the accuracy of prediction. Moreover, our new hybrid model appears to be superior in predicting U-SMEs' default events during the financial crisis and within a short-time span, a vital aspect for all banks engaging in U-SMEs' lending activities as part of their day-to-day operations.

The paper proceeds as follows: section 2 reviews the market-based default prediction models literature. Section 3 presents the data collection procedure, the estimation of the relevant modelling variables and introduces our new hybrid model. Section 4 discusses the empirical findings on its performance. Section 5 concludes.

2. Prior literature

2.1. Lending technologies

Traditional bank lending approaches are typically distinguished between transaction-based lending and relationship-based lending. As prior literature suggests, both technologies appear to have important benefits but their adoption is typically determined by factors such as the size of the borrowing firm, its financial history, the transparency of the borrower's financial statements, but also the size and organisational structure of the lender (Berger and Udell, 2002; Baas and Schrooten, 2006; Beck and Demirguc-Kunt, 2006; Hernández-Cánovas and Martínez-Solano, 2010; Bartoli et al., 2013). The former technology,

---

[8] Meaning that stock prices "fully reflect" all available information, i.e. company- and non-company-specific.
[9] As market prices are a function of anticipated future cash flows, wider economic events such as potential economic recessions or financial crises will be reflected in asset values via this pricing mechanism. Using L-SMEs market data for U-SMEs' default prediction allows us to take into account the impact of such events on the organisational viability of the latter firms on an ex ante basis.

transaction-based lending, encapsulates the three lending technologies of financial statement lending, asset-based lending, and credit scoring. This is typically used by larger banks utilising quantitative data obtained from the borrowers' financial statements backed with appropriate collateral guarantees (Berger and Udell, 2006).[10] In contrast, SME lending is predominately driven by the use of relationship lending technologies. These technologies require closer monitoring of the SME and allow access to qualitative information through frequent and personal interaction between the loan officer and the manager of the firm (Berger and Udell, 2002). Such technologies are often considered as a panacea for bank-SME relationships as they allow lenders to grant access into proprietary information of otherwise 'opaque' SMEs. Extant literature shows that relationship lending leads to the increased value of such information (Berger and Udell, 2002, 2006; Schæffer, 2003; Boot et al., 2005; D'Aurizio et al., 2015), reduction in information asymmetry (Berger et al., 1999; Boot, 2000), optimal lending decision-making for smaller banks (Berger and Black, 2011), lower borrowing costs for the SMEs (Peterson and Rajan, 1994; Schæffer, 2003) and continuation of credit lines for SMEs specially during financial crises (Bolton et al., 2016).

Nonetheless, not all studies are supportive to this view. Baas and Schrooten (2006) show that interest rates on loans to SMEs are unrelated to the duration of the lending relationship between the two parties and that borrowers are charged higher interest rates when banks rely on relationship lending technologies as opposed to those markets where both alternative lending technologies exist.[11] A handful of studies also show that both technologies (transaction-based and relationship-based) tend to be complementary to each other (Uchida et al., 2008; Muro, 2010; Bartoli et al. 2013). Bartoli et al. (2013) suggest that this complementarity of lending technologies is possibly driven by efficiency considerations such

---

[10] These lending technologies are also used, to some extent, for SMEs with long financial history and audited financial statements.
[11] This study corroborates the evidence produced by Petersen and Rajan (1994), Angelini et al. (1998) and Degryse and Van Cayseele (2000).

as the need of lenders to increase the degree of delegation (decentralisation) and lower the turnover of branch managers. They also show that soft information tends to be used in combination with hard information when lenders are using both technologies as a primary tool for lending decisions.

2.2. Approaches to default prediction

There is extensive literature spanning more than three decades on business failure prediction with the majority of such studies concentrating on the use of various models on large and publicly listed companies.[12] Regarding SMEs' default prediction, prior research is rather scarce with just a handful of studies simply suggesting the importance of developing more advanced models specific to small firms' characteristics (Altman and Sabato, 2007; Altman et al., 2010). Moreover, as the majority of such firms are typically unlisted, their evaluation in terms of credit risk is predominantly carried out (in practice) using accounting-based models. However, such models manifest significant theoretical and practical weaknesses. Firstly, reliance on accounting statements to assess credit worthiness tends to be unreliable given that accounting ratios are based on historical information and therefore not necessarily informative in predicting SMEs' future performance (Agarwal and Taffler, 2008).

This is more prevalent in the case of their assets where the use of historical cost accounting results in true asset values usually being very different to their recorded book values. As Hillegeist et al. (2004) argue, accounting-based models are designed under the assumption that firms are not likely to go bankrupt resulting in the asset value of the company, especially fixed and intangible assets, often being overestimated in the financial statements. This problem is further amplified by potential managerial incentives to manipulate accounting information, under conditions of financial distress, and is found to be

---

[12] For a more comprehensive review of the different approaches readers can refer to Duffie and Singleton (2003), Tudela and Young (2003), Charitou et al. (2004), Vassalou and Xing (2004), Balcaen and Ooghe (2006), Bharath and Shumway (2008) amongst others.

more extensive in the case of SMEs where owners tend to be also the company's managers (Campa and Camacho-Miñano, 2015).

Empirically, Richardson et al. (1998) show that accounting-based models perform significantly worse compared to the use of trained experts (loan officers) as these models are unable to control for information changes caused by business cycles and are unable to control for issues such as the volatility of the firm's assets. This large deterioration in predictability is more evident during recessionary periods where there is a rapid escalation in the number of bankruptcies (Bilderbeek and Pompe, 2005). This problem is more serious for the case of U-SMEs, as the likelihood of such companies failing during times of recession is considerably higher than average. As Shumway (2001) argues, using only accounting ratios in default prediction models leads to suboptimal decision-making as market-driven variables such as past stock returns, their variability and the firm's market size are all significantly related to default. The author goes further, proposing a model that produces out-of-sample forecasts using both accounting-ratios and market-driven variables that demonstrate an increased level of prediction accuracy compared to other alternatives. Subsequent empirical tests by Chava and Jarrow (2004) confirm the superior forecasting performance of Shumway's (2001) model over previous modelling approaches such as those of Altman (1968) and Zmijewski (1984).

A significant development in solving the aforementioned problems and in improving bankruptcy prediction was the introduction of structural models that utilise option pricing theory in corporate debt valuation such as those introduced by Black and Scholes (1973) and Merton (1974). In particular, the latter approach by Merton (1974), *Merton DD model thereafter*, provided the foundation for all subsequent market-based default prediction models currently present in the literature. Its advantage is the provision of an intuitive picture as well as an endogenous explanation for credit default by connecting elements of credit risk to underlying structural variables and incorporating option pricing methods in default

prediction. As Wang (2009) argues, this model, including subsequent variants, not only facilitates security valuation but also addresses the choice of alternative capital structures.

Prior literature overwhelmingly suggests that the use of the Merton DD model can accommodate most of the aforementioned criticisms of accounting-based models as it provides a methodological approach that is unlikely to be affected by a firm's accounting policies and is not time- or sample-dependent (Hillegeist et al., 2004; Reisz and Perlich, 2004; Vassalou and Xing, 2004; Bharath and Shumway, 2008; Campbell et al., 2008). However, as Agarwal and Taffler (2008) argue, neither accounting-based models nor market-based models are exclusively sufficient for failure prediction as both incorporate unique sets of company information.

Empirical testing of the Merton DD model in the UK confirms that its distance-to-default measure (DD) is the most significant variable for measuring credit risk. With regard to the use of accounting variables, these appear to be incrementally informative when added to the main model (Demirovic and Thomas, 2011). This finding corroborates prior US literature on the usefulness of combining market-based and accounting-based information in predicting firm's default (Beaver et al., 2005; Campbell et al., 2008). Benos and Papanastasopoulos (2006) take a step further in modelling corporate default by introducing a hybrid model derived from the combination of various credit risk approaches. Their study uses an ordered probit regression model where the neutral distance to default is estimated from a series of financial ratios while accounting-based measures are utilised as explanatory variables.[13] This new hybrid model demonstrates improved in-sample fitting credit ratings and out-of-sample default predictability. These findings are also corroborated by Bellalah et al. (2016) and Doumpos et al. (2014) who examine default risk predictability for French companies and European listed firms respectively.

---

[13]In order to examine the efficiency of their hybrid modelling approach, Benos and Papanastasopoulos (2006) estimate two different models in which risk neutral distance-to-default metric and financial ratios are used separately as indicators of the firm's default.

Regarding the use of the Merton DD model in SMEs, there is only a handful of prior studies due to the lack of available market information for such firms. Using a sample of 246 L-SMEs in the UK for the period 2001-2004, Lin et al. (2007) report higher predictive ability of the Merton DD model for the short run (<1 year); while, accounting-based models are found to be superior, in terms of accuracy, in the long run (>1 year). Finally, Chen et al. (2010) use a similar approach in the examination of Chinese L-SMEs and suggest that the predictive accuracy of the adjusted Merton DD model is highly sensitive to the identification of the various default points.

Nonetheless, although the use of the above modelling approach demonstrates good performance in terms of predicting SME default, it has only been applied to L-SMEs. Unfortunately, these firms (L-SMEs) constitute only a very small proportion of the entire SME sector in the UK economy. This is also the case for the market-based methodological approaches used in the most recent studies of Doumpos et al. (2014) and Bellalah et al. (2016) where, unlike listed firms, U-SMEs' distance-to-default information is not available. Given the points discussed earlier regarding the reliability of accounting-based models and the importance of market-based information in predicting default events, it is of great academic and practitioner interest to develop a hybrid default prediction model that combines market and accounting information to predict the default events in the case for U-SMEs. We now proceed to the discussion of our modelling approach that attempts to solve this problem and fill this crucial gap in the literature.

3. Data and methodology

3.1. Data

Our study employs a sample of listed (L-SMEs) and unlisted SMEs (U-SMEs). All L-SMEs are selected based on the Basel definition for small firms, i.e. firms with total turnover

value of less than £42 million. All operating and financial indicators are obtained from Thomson's Datastream. The final sample of L-SMEs is 198 companies, all of them in the manufacturing sector covering the period from 2004 to 2013. With regards to the UK U-SMEs' sample, our dataset consists of approximately 20,000 companies per annum with all relevant financial information obtained from Bureau Van Dijk's FAME database. This sample covers both active (non-defaulted) and dead (defaulted) SMEs with the latter category including all firms in liquidation, administration and receivership during the period under examination.

According to Table 1, the total number of defaults observed in our U-SMEs sample based on aforementioned criteria is 14,170 out of a total sample of 196,807 SME observations. This sample is unbalanced with the number of firms changing every year due to various corporate events such as bankruptcy and M&A activities. Similarly, the default rates in our sample also vary from year to year with a reported increase during the credit crunch period of 2007-09. For instance, the default rates of U-SMEs for 2007, 2008 and 2009 are 7.21%, 8.56% and 9.44% respectively while the default rate for the entire period under examination is 6.72%. A similar trend is observable in the case of L-SMEs with 5.79% for 2007, 6.42% for 2008 and 6.49% for 2009. However, as Table 1 shows, the total default rate for L-SMEs compared to U-SMEs is lower (5.54% vis-à-vis 6.72%). This is not an unexpected finding given that L-SMEs have better access to capital and inevitably are more able to overcome possible financial difficulties.

[Insert Table 1 about here]

Table 2 reports the statistical properties of our L-SMEs and U-SMEs sample using a number of accounting indicators.[14] All variables are winsorized 5% in each tail to reduce the impact of outliers. Although the average retained earnings to total assets ratio (*RETA*) for U-

---

[14] All accounting ratios used in our study are in line with prior literature in credit default prediction (Altman and Sabato, 2007; Altman et al., 2010).

11

SMEs and L-SMEs is -0.2534 and -0.0103 respectively, U-SMEs appear to be more profitable than their listed counterparts with an average *EBITTA* of 0.6155 (0.2479 for the case of L-SMEs) and an average net income to sales ratio (*NIS*) of 0.3245 (0.2277 for L-SMEs). Moreover, the average short-term leverage (*LEV*) for U-SMEs is 1.5666 compared to 1.5072 for L-SMEs indicating that the former type of firms are marginally more reliant on short-term borrowing despite the fact that they generally benefit from higher profitability. U-SMEs also tend to hold more cash within their asset structure with an average cash to total assets ratio (*LIQ*) of 0.1469 as compared to 0.0639 for L-SMEs. This cautionary approach in financial management is more likely to be attributed to the U-SMEs' greater reliance on short-term borrowing as opposed to long-term borrowing which is indicated by the current liabilities to non-current liabilities ratio (*CLNCL*) of 19.035 (14.621 for the L-SMEs). Overall, the picture emerging from these statistics is that U-SMEs are, on average, more profitable and operationally efficient but also more prone to face short-term liquidity problems (financial distress). For example, although the average performance in terms of cash to net sales (*CNS*) and net cash to net worth (*NCNW*) is 0.1058 and 0.3229 (0.0722 and 0.1351 for their L-SME counterparts), their performance in terms of working capital to total assets (*WCTA*) and current assets to current liabilities (*CACL*) ratios is significantly lower with 0.0929 and 1.6891 as compared to 0.1350 and 2.0423 in the case of the L-SMEs.

[Insert Table 2 about here]

Table 3 also gives an insight on the absolute size of L-SMEs and U-SMEs in our sample by comparing the average profitability, sales, total assets and total liabilities. L-SMEs have, on average, higher total liabilities (£5.437m) and total assets (£7.755m), while sales and profitability are, on average, also substantially higher with £35.769m and £7.696m compared

to their non-listed counterparts (£15.353m and £1.964m respectively). [15] These results indicate that L-SMEs are on average larger firms, less geared and more efficient in generating sales and profits for their equity holders. This finding is further supported by the reported average median values.

[Insert Table 3 about here]

3.2. Merton-KMV modelling approach

Our market-based indicators of default are estimated from the sample of UK L-SMEs followed by the use of appropriate calibration techniques needed to obtain the hybrid indicators of default that combines both accounting and market information for the full sample. These calibration techniques follow the Merton-KMV approach which is an established structural credit risk model used to test the effect of market-based information on corporate default (Crosbie and Bon, 2003; Chen et al., 2010). This approach is a modified version of the Merton (1974) framework and attempts to predict firms' default based on their underlying debt structure.

Based on the original Merton DD model, all payoffs to the shareholders of the firm are similar to the payoffs from the call option on the firm's assets with debt outstanding being the exercise price. Hence, the model assumes that firms should have a single issue of zero-coupon debt outstanding *(D)* which means that at a specified maturity date (τ) an amount of *D* is due. As a rule of thumb, at maturity date (τ), the face value of debt (*D*) would be received by debt holders if there is enough asset value *(V)* to meet this payment. Hence, at maturity date if *V>D*, debt holders would receive *D* and equity holders will get the rest (*V-D)*. However, if the value of the firm's assets is not sufficient to satisfy the debt holders claims (*V<D),* debt holders will receive the total value of the firm's asset and equity holders

---

[15] Although our methodological approach does not control for size, as it is based on aggregate sector-level data, we do believe that future improvements in the approach can be easily made by utilising a multi-attribute matching procedure.

will receive nothing. Our estimation procedure starts by considering the equity value as an option on the value of the firm. It is noteworthy that in this case, on date τ, equity holders receive any amount remaining after the debt holders are paid off mathematically notated as

$$\begin{cases} V - D, & if \quad V > D \\ 0, & otherwise \end{cases} \tag{1}$$

The payoff to equity holders expressed above is replicating the payoffs of a long call option on the firm's value with maturity τ and strike price D. Hence the call value is equal to the value of equity.

The main problem in implementing the Merton DD model is that the firm's asset value V and its volatility (σ_V), two essential elements for estimating the distance to default *(DD)*, are both unobservable.[16] Unlike the value of equity *(E)* and equity volatility *(σ_E)*, both of which can be easily proxied by the use of market capitalisation (Das and Sundaram, 2004), V and σ_V have to be inferred. Crosbie and Bohn (2003) mention that the relationship between equity volatility *(σ_E)* and a firm's asset volatility *(σ_V)* indicated in the Merton DD model might not lead to reasonable results as market leverage varies considerably in practice leading to a biased estimation of the probability of default.[17] To overcome this problem, we adopt the Merton-KMV approach and use the Newton-iterative procedure to calculate V and σ_V as our unknown parameters.[18] The average number of iterations needed for each L-SME in our sample to reach convergence is 3, while all cases where a convergence criterion is not

---

[16] If the firm's asset value and its volatility are available, the firm's probability of default could be estimated easily. For example, other relevant studies such as that of Bharath and Shumway (2008) and Charitou et al. (2013) estimate the firm's value as the sum of the firm's equity value and debt value.

[17] For instance, when a firm's credit risk is improving, the model might overstate the probability of a firm's default because the asset volatility will be overestimated if the market leverage goes down quickly. On the other hand, at a time of rapid increase in market leverage, the firm's asset volatility will be underestimated. In this situation, in spite of the deterioration in the firm's credit risk, the model understates the probability of the firm's default.

[18] Convergence typically exists when the difference between the newly estimated asset value volatility (σ_V') and the true asset value volatility (σ_V) is less than 10^{-5}. A detailed explanation of the mathematical process for the Merton DD model is provided in Bharath and Shumway (2008) and for the Newton-iterative method of the Merton-KMV approach in Chen et al. (2010).

reached are eliminated from our estimation sample ($n = 24$). All L-SMEs' equity values *(E)* are extracted from Datastream for each calendar year-end. Likewise, equity value volatility (σ_E) is proxied using the standard deviation of the sample firms' daily equity value *(E)* for each year during the 2004-2012 sample period.[19]

A crucial parameter in the Merton-KMV model is the default point (*DPT*). According to Crosbie and Bohn (2003), firms generally remain solvent when their asset value is up to the book value of their total liabilities. However, at the time of default, the value of the firm's assets is commonly between the value of current liabilities and that of the total liabilities. This is algebraically formulated as:

$$DPT = CL + k \times LL, \quad 1 \geq k \geq 0, \tag{2}$$

where, *DPT* denotes the default point, *CL* stands for current liabilities and *LL* denotes the long-term liabilities of the firm. Prior literature suggests that the predictive accuracy of the model is sensitive to the default point changes (Huang and He, 2010; Lee, 2011). To ensure comparability with prior studies we estimate the default point in our model using *k = 0.5*. The last two integral parameters for building up the Merton-KMV credit default model are those of liability maturity (τ) and the risk free rate (r). This study uses the one-year liability maturity for τ , and the Bank of England one-year base rate for r.

Having obtained V and σ_V , we can then calculate the distance to default *(DD)* for the remaining 174 L-SMEs from our initial sample (87.8%) defined as:

$$DD = \frac{V - D}{V \sigma_V} \tag{3}$$

where V is the firm's asset value, D are the debts in default points, and σ_V is the asset value volatility. Moreover, the corresponding implied probability to default for each of the L-

---

[19] Daily equity values are estimated using daily closing stock prices × number of outstanding shares over 252 trading days. Equity volatility is also estimated annually. This method is in line with prior studies in the field (Hull et al., 2004).

SMEs, often reported as expected default frequency (*EDF*) in the literature, is calculated as

$$EDF = N(-DD) \tag{4}$$

where *N* is the cumulative standard normal distribution and *DD* is the distance to default.

3.3. Estimating hybrid indicators of default for the U-SMEs' sample

To address the problem of estimating default risk for SMEs that are not publicly listed, we introduce an innovative approach that allows us to capture the effect of both accounting and market information by combining the Merton-KMV process with the traditional use of a logistic regression model. The process is implemented in four steps which are presented in Figure 1.

[Insert Figure 1 about here]

In detail, we initially estimate the Merton-KMV probability of default (EDF) for the 174 L-SMEs and for each year during our sample period 2004 to 2012 using the process described in section 3.2. The next step is to estimate individual L-SMEs' default scores (X_t^{LSME}). This is accomplished by using the inverse of the logistic function, defined as

$$X_{it}^{LSME} = \ln \frac{KPG_{it}^{LSME}}{(1 - KPG_{it}^{LSME})} \tag{5}$$

where, X_{it}^{LSME} represents individual L-SMEs' logistic scores at time *t,* and KPG_{it}^{LSME} is the individual probabilities of an L-SME not defaulting at time *t,* estimated as $(1 - EDF_{it}^{LSME})$.[20]

Once all individual default scores (X_{it}^{LSME}) are calculated, we are then using them as dependent variables in a linear regression model with L-SMEs' accounting indicators being the explanatory variables using a forward stepwise selection procedure in line with Altman and Sabato (2007). Our aim at this stage is to generate the relevant market-based coefficients

---

[20]As Altman and Sabato (2007) argue, the use of the *Known Probability of Being Good* (KPG) is superior as it allows us to have positive slopes and positive intercepts given that higher logit scores indicate a lower probability that a firm will default.

that will be used to estimate the default probabilities for the U-SMEs sample. Hence, our linear regression model is formulated as follows:

$$
\begin{aligned}
X_{it}^{LSME} = \ &\beta_0 + \beta_1 RETA_{it}^{LSME} + \beta_2 STA_{it}^{LSME} + \beta_3 COV_{it}^{LSME} + \beta_4 LEV_{it}^{LSME} + \beta_5 LIQ_{it}^{LSME} \\
&+ \beta_6 EBITTA_{it}^{LSME} + \beta_7 NIS_{it}^{LSME} + \beta_8 TLTA_{it}^{LSME} + \beta_9 CACL_{it}^{LSME} + \beta_{10} CNS_{it}^{LSME} \\
&+ \beta_{11} CETL_{it}^{LSME} + \beta_{12} NCNW_{it}^{LSME} + \beta_{13} STDNW_{it}^{LSME} + \beta_{14} CLNCL_{it}^{LSME} + \beta_{15} WCTA_{it}^{LSME} + \varepsilon_t
\end{aligned}
\tag{6}
$$

where, X_{it}^{LSME} is the individual logistic scores for the L-SMEs at time t described in Eq.(5), *RETA* is the ratio of retained earnings to total assets used as a proxy for profitability; *STA* is the ratio of sales to total assets (activity); *COV* is the ratio of EBITDA to interest expenses (debt coverage); *LEV* is the ratio of short-term debt to equity book value (leverage); *LIQ* is the ratio of cash to total assets (liquidity); *EBITTA* is the ratio of earnings before interest and tax to total assets (profitability); *NIS* is the ratio of net income to net sales (profitability); *TLTA* is the ratio of total liabilities to total assets (leverage); *CACL* is the ratio of current assets to current liabilities (liquidity); *CNS* is the ratio of cash to net sales (efficiency); *CETL* is the ratio of total capital employed to the total liabilities (leverage); *NCNW* is the ratio of net cash to net worth (liquidity); *STDNW* is the ratio of short-term debt to net worth (financial distress); *CLNCL* is the ratio of current liabilities to the non-current liabilities (debt structure); *WCTA* is the ratio of working capital to total assets (liquidity) and, ε_t is the error term.

The last stage of our calibration method involves the estimation of our 'hybrid'/marker-based default score ($KPG_{it}^{mb,USME}$) for all U-SMEs. This is done by employing a formula that combines the market-based coefficients from Eq. (6) alongside the accounting indicators for each individual unlisted firm in our 2004-2012 sample mathematically defined as

$$
\begin{aligned}
KPG_{it}^{mb,USME} = \ &\beta_0 + \beta_1 RETA_{it}^{USME} + \beta_2 STA_{it}^{USME} + \beta_3 COV_{it}^{USME} + \beta_4 LEV_{it}^{USME} + \beta_5 LIQ_{it}^{USME} \\
&+ \beta_6 EBITTA_{it}^{USME} + \beta_7 NIS_{it}^{USME} + \beta_8 TLTA_{it}^{USME} + \beta_9 CACL_{it}^{USME} + \beta_{10} CNS_{it}^{USME} \\
&+ \beta_{11} CETL_{it}^{USME} + \beta_{12} NCNW_{it}^{USME} + \beta_{13} STDNW_{it}^{USME} + \beta_{14} CLNCL_{it}^{USME} + \beta_{15} WCTA_{it}^{USME}
\end{aligned}
\tag{7}
$$

where, *mb* stands for market-based, and all other variables are described above.

3.4. Estimation of accounting-based indicators of default for U-SMEs' sample

To evaluate the performance of our model, we compare its accuracy with that of an existing accounting-based one. Prior literature suggests that the use of accounting-based models can be appropriate in predicting SMEs' default (Altman and Sabato, 2007; Altman et al., 2010). Comparing the performance of our hybrid market-based model to the plain accounting-based model allows us to assess the suitability and potentially superior efficiency of our modelling approach. The plain accounting-based model is a standard logistic regression that uses the same accounting ratios as above. Our accounting-based model is defined as:

$$
\begin{aligned}
KPG_{it}^{ab,USME} = {} & \beta_0 + \beta_1 RETA_{it}^{USME} + \beta_2 STA_{it}^{USME} + \beta_3 COV_{it}^{USME} + \beta_4 LEV_{it}^{USME} + \beta_5 LIQ_{it}^{USME} \\
& + \beta_6 EBITTA_{it}^{USME} + \beta_7 NIS_{it}^{USME} + \beta_8 TLTA_{it}^{USME} + \beta_9 CACL_{it}^{USME} + \beta_{10} CNS_{it}^{USME} \\
& + \beta_{11} CETL_{it}^{USME} + \beta_{12} NCNW_{it}^{USME} + \beta_{13} STDNW_{it}^{USME} + \beta_{14} CLNCL_{it}^{USME} + \beta_{15} WCTA_{it}^{USME} + \varepsilon
\end{aligned}
\tag{8}
$$

where, $KPG_{it}^{ab,USME}$ stands for Known Probability of Being Good for each U-SME at time *t,* and *ab* stands for accounting-based. In line with Altman and Sabato (2007), this is a binary variable with all U-SMEs that report their status as "active" being assigned the score 1, while firms with company status "in liquidation", "in administration" and "in receivership" are given the value of 0. Finally, to account for the possible timing effects of default prediction accuracy and reliability all empirical models described above include relevant time controls.[21]

4. Results

4.1. Default prediction for the L-SMEs sample using the Merton-KMV model.

As it is assumed in Merton-KMV model, the asset value is subject to normal distribution and the default distance reflects the standard deviation from the company's default. Thus,

---

[21] One dummy from each set was dropped to avoid multicollinearity.

firms' expected default probability (EDF^{LSME}) can be calculated using the normality function of distance to default (*DD*). According to Table 4, the average EDF^{LSME} value for the sample is approximately 1.7 percent (1.695) while the minimum value is 0.000 and the maximum is 27.847 percent. By observing individual EDFs, we notice that a large number of default probabilities within the sample (78%) range between 0.0 and 2.0 percent. However, there are very few observations (outliers) for which the EDF^{LSME} score is substantially greater than zero indicating that some L-SMEs have a high probability of default.

[Insert Table 4 about here]

In order to evaluate the performance of the market-based model we test its classification accuracy by obtaining the Receiver Operating Characteristic (*ROC*) curve for the L-SMEs within our 2004-2012 sample period. According to Figure 2, the area under the ROC curve (*AUC*)[22] is 0.8754 which indicates the performance of the Merton-KMV model in predicting L-SMEs' default events. This model is implemented as part of the calibration process to derive the 'hybrid' coefficients that will subsequently be used in the estimation of default scores for the entire U-SMEs sample. The accuracy of our model (AUC of 0.8754) appears to be slightly better than that of a similar model used in the case of Chinese SMEs (AUC of 0.85) and the same *k=0.5* default point (Chen et al., 2010).

[Insert Figure 2 about here]

4.2. Market-based (hybrid) and accounting-based model results for the L-SMEs sample

To avoid possible bias into our statistical estimates that might be introduced by outliers in our data, all variables are winsorized 5% in each tail for both cases, i.e. the 'hybrid' and the accounting-based regression models. Results on the use of the 'hybrid' logistic scores for

---

[22] The area under the ROC curve and the equivalent index, the Gini Coefficient, are widely used to measure the performance of classification models. The AUC is a measure of the difference between the score distributions of failed and non-failed companies and the Gini coefficient is an index which can be calculated as $[(2 \times AUC) - 1]$.

L-SMEs (X_{it}^{LSME}) is presented in Table 5. According to our findings, the only insignificant variable in the model is the ratio between current and non-current liabilities (*CLNCL*) for which the coefficient is approximately zero. All other explanatory variables in this model are mostly highly significant predictors in statistical terms of the 'hybrid' X-score for our L-SMEs sample. For example, all profitability indicators, *RETA, EBITTA* and *NIS* show a statistically significant positive relationship with their X_{it}^{LSME} logistic scores which is in line with standard finance theory, *i.e.* an increase in firms' profitability leads to an increase in the probability of not defaulting. The results on the leverage and liquidity indicators appear to be similar, also showing a positive relationship which is significant at the 1% level. Rather surprising, in terms of short-term liquidity, L-SMEs appear to be marginally influenced by an increase in short-term debt. For example, as the *LEV* variable show, although an increase in short-term liabilities has a negative impact on the probability of survival, the relevant coefficient is significant only at the 10% level. The overall picture emerging is that the L-SMEs, on average, tend to avoid large levels of debt in their capital structure, especially non-current ones, for which the payback and interest extends far into the future. Finally, the adjusted R-square of the linear regression is 85.45 percent and the F-test is also highly significant at the 1% level, showing that our model is well fitted.

[Insert Table 5 about here]

We are now comparing the performance of our hybrid model with the standard logistic accounting-based model. The results of this model regression are presented in Table 6.

[Insert Table 6 about here]

According to these results, all coefficient signs and significance levels are consistent to prior literature (Altman et. al, 2010). The only statistically insignificant variable in the model is *LEV* (coefficient of -0.0025) indicating no relationship between short-term debt increase

and probability of survival for the case of the U-SMEs.[23]

4.3. Performance test of the models on the U-SME sample

The performance of the model with market indicators is tested by comparing its ROC (AUC) with that of the standard accounting-based model in the 2013 hold-out sample and also within the entire sample period 2004 to 2012.[24] In order to create the hold-out sample, 166 defaulted firms were collected from the year 2013, while 2,265 active firms are also randomly selected to keep the default rate of the sample for UK U-SMEs equal to the default rate of the entire U-SMEs' sample for 2004-2012 (6.72%). We retained data from 2013 in order to undertake hold-out tests for model performance. The test is implemented by predicting the defaulted firms in the 2013 hold-out sample using the hybrid $KPG_{it}^{mb,USME}$ and the accounting-based $KPG_{it}^{ab,USME}$ logistic scores for each firm in the sample. The descriptive statistics of both default scores in the 2013 hold-out sample are reported in Table 7. The results indicate higher mean values, on average, for the former model compared to the latter. The mean value of the $KPG_{it}^{mb,USME}$ model score is 4.521, while for the $KPG_{it}^{ab,USME}$ model is only 2.574. Moreover, the standard deviation for the former model is higher, 1.172 and 0.732 respectively.

[Insert Table 7 about here]

These logistic scores are then used to estimate the probabilities-to-default (PDs) for both models in the 2013 hold-out sample and the full 2004-12 periods, while the predictive ability of both models is tested using the ROC reports. The results of ROC (AUC) indicate that the

---

[23]A careful examination of the test statistics for the accounting model in Table 6 reveals minor specification issues, i.e. log-likelihood of -45720 and Wald Chi-square 3783 significant at the 1% level. As this logistic regression replicates models used previously in the literature (Altman and Sabato, 2007; Altman et al., 2010) we still report the results only for comparability purposes. We leave further improvements on this modelling approach for future research.
[24]This method has been used as a validation technique by many relevant studies such as Altman et al. (2010) and Chen et al. (2010) to test the accuracy of their models.

model with the market information incorporated on it performs better when compared to the plain accounting-based one. According to Figure 3, the area under the ROC curve for the hybrid model in the 2013 hold-out sample is 0.8387 which indicates a good level of classification accuracy. This is evidently superior to the ROC (AUC 0.7901) for the model based solemnly on accounting information. Similarly, we report the ROC tests for both models within the entire 2004-2012 sample. The results from the within-sample test further confirm the superiority of the hybrid model in demonstrating better performance over its pure accounting-based counterpart (AUC-values of 0.8479 and 0.8114 respectively). These results are also superior to the results of Altman et al. (2010) in which the UK SME model's accuracy for the hold-out sample is lower (AUC 0.76 and AUC 0.75) than that of our modelling approach (AUC 0.8387). Moreover, our model's overall accuracy (AUC 0.8479) is also higher than that from Altman et al. (2010) who report AUC of 0.78 and 0.80 within the entire sample. This clearly indicates the importance of using market-wide information in predicting U-SMEs' default events and the merits of our innovative modelling approach.

[Insert Figure 3 about here]

One of the main issues with the use of accounting-based default prediction models discussed in the literature is that they are not reliable during periods of financial crises. This issue can be more problematic for the case of U-SMEs as accounting-based models and scorecards are widely used by all credit providers for predicting SMEs' default. On the one hand, as the likelihood of SMEs' financial failure during recessions is higher, it is essential for the banks to be able to use the default prediction model with the highest predictive power to avoid potential losses. On the other hand, the SMEs' accounting-based default prediction models are found not to be accurate when used for short sample periods. This problem becomes more severe during periods of financial turmoil when it is essential for all credit providers to predict the firm's default within short-term time frames, given that financially weak firms

22

might not manage to survive such troublesome periods.

We implement a test to investigate the predictive performance of our hybrid model against the accounting-based one during the financial crises and also for short-term periods. To do so, a hold-out sample is created in the end of the financial year 2008 with 150 defaulted firms and 2,056 active firms from our sample.[25] Using the hybrid and the accounting-based models, we estimate the relevant *KPGs* with the aim to predict the SMEs' default events in the 2008 hold-out sample. This allows us to test the accuracy of the two models by observing the company status at the end of 2009 and comparing that to the relevant predicted defaults. The results of the accuracy tests using the ROC curve are illustrated in Figure 4.

[Insert Figure 4 about here]

As we can see from Figure 4, the accuracy of the hybrid model ($KPG_{it}^{mb,USME}$) in predicting U-SMEs default during the financial crisis is far superior with an AUC-value of 0.8362 compared to the AUC-value of 0.7781 for the accounting-based model ($KPG_{it}^{ab,USME}$). Further tests on the predictive accuracy of our model for U-SMEs' default within short-time periods produce a similar picture. By comparing the estimated *KPGs* from the financial years 2004-2012, we observe again that the accuracy of the hybrid model within annual intervals is superior to that of the standard accounting-based model for 8 out of the 9 years in our sample. According to Table 8, it is only for 2011 where the prediction accuracy of the hybrid model (79.81%) is marginally lower than the one of the accounting-based model (80.31%).

[Insert Table 8 about here]

The accuracy rate of the hybrid model for the rest of our sample period appears to be considerably better. Hence, based on these results we can conclude with confidence that our proposed market-based model performs better in short-time periods which is very critical for

---

[25] The default rate of the 2008 hold-out sample is also equal to the default rate of the entire SME sample (6.72%).

all SMEs and especially for the case of unlisted ones where their financial performance is less monitored by relevant market forces and creditors alike.

5. Conclusion

An accurate U-SMEs' default prediction model is crucial for all relevant stakeholders, including banks and other credit providers, the companies' owners, government agencies, accounting professionals, etc. It primarily helps banks to assess U-SMEs' financial prospects with a higher level of accuracy, reducing the lenders' potential losses due to credit misallocation. It also enables banks to estimate the capital requirement for their U-SMEs' lending portfolio more accurately by truly reflecting the riskiness of such firms. As a more accurate prediction model reduces the chance of credit misallocation, the funds will be distributed fairly among the entire SMEs' lending portfolio which can lead to corporate growth and to minimisation of lost income from the banks' perspective.

The comparative results of our study indicate that the hybrid U-SMEs' default prediction model which employs market-based information along with the accounting-based information performs considerably better when compared to its accounting-based counterpart for (i) the entire period under investigation, (ii) during and after the period of the financial crisis; as well as (iii) during short-term default prediction time frames. These results make our modelling approach an elaborate default prediction tool for such types of firms, where any monitoring mechanisms such as those typically imposed by market forces in the case of publicly-listed firms is intrinsically absent. This is because both accounting and market information appear to be incrementally informative to each other when assessing the credit quality of a firm. We argue that the methodological approach of including market information of listed-SMEs along with accounting information of U-SMEs is therefore not only crucial but also a very reliable mechanism in predicting default for such firms and should be used

confidently by all finance providers.

In this respect, we believe that the proposed modelling approach is not only beneficial to those banks advocating transactional lending technologies but also those that employ relationship-based ones. More recent literature shows that during financial crises relationship banking ensures better continuation of credit facilities mostly to profitable firms, especially in the case of U-SMEs (D'Aurizio et al., 2015; Bolton et al., 2016). As the method proposed in this paper increases the accuracy of prediction during financial crises for U-SMEs, it can provide banks advocating hard-information technologies with a useful tool to safeguard their assets during aggregate credit contractions without having to restrict lending across their entire U-SMEs' portfolio.

Furthermore, its ability to create a basis for lenders to compare between listed- and unlisted-SMEs makes it a suitable and easy-to-use initial screening tool alongside typical credit scoring techniques before the gathering of soft information with the use of loan officers takes place. This may have important implications for both the cost of the lending service to smaller credit providers but also to the larger banks that are predominately engaged in hard information lending.

References

Agarwal, V., & Taffler, R. (2008). Comparing the performance of market-based and accounting-based bankruptcy prediction models. *Journal of Banking and Finance,* 8, 1541-1551.

Alford, A.W. (1992). The effect of the set of comparable firms on the accuracy of the price-earnings valuation method. *Journal of Accounting Research*, 30(1), 94-108.

Altman, E.I. (1968).Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 23(4), 589-609.

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

Altman, E.I., Sabato, G., & Wilson, N. (2010).The value of non-financial information in small and medium-sized enterprise risk management. *Journal of Credit Risk*, 6(2), 95-127.

Angelini, P., Di Salvo, R., & Ferri, G. (1998). Availability and cost of credit for small businesses: customer relationships and credit cooperatives. *Journal of Banking and Finance,* 22, 925–954.

Baas, T., & Schrooten, M. (2006). Relationship banking and SMEs: A theoretical analysis. *Small Business Economics*, 27(2), 127-137.

Baker, M., & Ruback, R. (1999). Estimating industry multiples. *Harvard University, Working paper*

Balcaen, S. & Ooghe, H. (2006). 35 Years of Studies on business Failure: An Overview of the Classic Statistical Methodologies and Their Related Problems. *The British Accounting Review.* 38(1), 63-93.

Bartoli, F., Ferri, G., Murro, P., & Rotondi, Z. (2013). SME financing and the choice of lending technology in Italy: Complementarity or substitutability? *Journal of Banking and Finance*, 37(12), 5476-5485.

Beaver, W.H., McNichols, M.F., & Rhie, J. (2005). Have financial statements become more informative? evidence from the ability of financial ratios to predict bankruptcy. *Review of Accounting Studies*, 10, 93–122.

Beck, T., & Demirguc-Kunt, A. (2006). Small and medium-size enterprises: Access to finance as a growth constraint. *Journal of Banking and Finance*, 30(11), 2931-2943.

Bharath, S.T., & Shumway, T. (2008). Forecasting default with the Merton distance to default model. *The Review of Financial Studies*, 21(3), 1339-1369.

Bhojraj, S., & Lee, C. (2002). Who is my peer? A valuation-based approach to the selection of comparable firms. *Journal of Accounting Research*, *40*(2), 407-439.

Bellalah, M., Zouari, S., & Levyne, O. (2016), The performance of hybrid models in the assessment of default risk. *Economic Modelling*, 52, 259-265.

Benos, A., & Papanastasopoulos, G. (2007). Extending the Merton Model: A Hybrid Approach to Assessing Credit Quality. *Mathematical and Computer Modelling*, 46(1-2), 47-68.

Berger, A.N. and Udell, G.F. (2002). Small business credit availability and relationship lending: the importance of bank organisational structure. *Economic Journal*, 112(477), 32–53.

Berger, A.N., Udell, G.F. (2006). A more complete conceptual framework for SME finance. *Journal of Banking and Finance*, 30, 2945–2966.

Berger, A.N., & Black, L.K. (2011). Bank size, lending technologies, and small business finance. *Journal of Banking and Finance*, 35(3), 724-735.

Bharath, S.T., & Sumway, T. (2008). Forecasting Default with the Merton Distance to Default Model, *The Review of Financial Studies*, 21(3), 1339-1369.

Bilderbeek, J., & Pompe, P.P.M. (2005).The Prediction of Bankruptcy of Small- and Medium-Sized Industrial Firms. *Journal of Business Venturing*, 20(6), 847-868.

BIS Economics Paper (2012). SMEs Access to External Finance, *Department for Business Innovation & Skills*.16.

Black, F., & Scholes, M. (1973).The pricing of options and corporate liabilities.*The Journal of Political Economy*, 81(3), 637–659.

Bolton, P., Freixas, X., Gambacorta, L., & Mistrulli, P.E. (2016). Relationship and transaction lending in a crisis. *The Review of Financial Studies*, 29(10), 2643-2676.

Boot, A.W. (2000). Relationship banking: what do we know? *Journal of Financial Intermediation*, 9(1), 7–25.

Boot, A.W., Milbourn, T.T., & Schmeits, A. (2005). Credit ratings as coordination mechanisms. *The Review of Financial Studies*, 19(1), 81-118.

Campa, D. & Camacho-Miñano, M. (2015). The Impact of SME's Pre-Bankruptcy Financial Distress on Earnings Management Tools. *International Review of Financial Analysis*, 42 (December), 222-234.

Campbell, J.Y., Hilscher, T., & Szilagyi, Y. (2008). In search of distress risk. *Journal of Finance*, 63, 2899–2939.

Charitou, A., Dionysiou, D., Lambertides, N., & Trigeorgis, L. (2013). Alternative bankruptcy prediction models using option-pricing theory. *Journal of Banking and Finance*, 37, 2329-2341.

Chava, S., & Jarrow, R. (2004). Bankruptcy Prediction with Industry Effects. *Review of Finance*, 8, 537-569.

Chen, X., Wang, X., & Wu, D. (2010). Credit Risk Measurement and Early Warning of SMEs: An Empirical Study of Listed SMEs in China. *Decision Support System*, 49, 301-310.

Claessens, S., Tong, H., & Wei, S. J. (2012). From the financial crisis to the real economy: Using firm-level data to identify transmission channels. *Journal of International Economics*, *88*(2), 375-387.

Competition and Markets Authority (CMA) & Financial Conduct Authority (FCA) (2014). *Banking services to small and medium sized enterprises. A CMA and FCA market study.* London: CMA/FCA

Crosbie, P., & Bohn, J. (2003). Modelling default risk. *Moody's KMV Company*. Available at https://business.illinois.edu/gpennacc/MoodysKMV.pdf.

D'Aurizio, L., Oliviero, T., & Romano, L. (2015). Family firms, soft information and bank lending in a financial crisis. *Journal of Corporate Finance*, 33, 279-292.

Das, S.R., & Sundaram, R.K. (2004). A Simple Model for Pricing Securities with Equity, Interest-Rate and Default Risk.*NYU Working Paper*.

Degryse, H., & Van Cayseele, P. (2000). Relationship lending within a bank-based system: evidence from European small business data. *Journal of Financial Intermediation* 9, 90–109.

Demirovic, A., & Thomas, D.C. (2011). The Relevance of Accounting Data in the Measurement of Credit Risk.*The European Journal of Finance*, 13(3).

Doumpos, M., Niklis, D., Zopounidis, C., & Andriosopoulos, K. (2014). Combining accounting data and a structural model for predicting credit ratings: Empirical evidence from European listed firms. *Journal of Banking & Finance,* 50,599-607.

Duffie, D., & Singleton, K.J. (2003).Credit Risk: Pricing, Measurement, and Management. *Princeton University Press*.

Dullmann, K., & Koziol, P. (2013). Evaluation of minimum capital requirements for bank loans to SMEs. *Deutsche Bundesbank*, 22.

Guariglia, A., Spaliara, M., & Tsoukas, S. (2016). To What Extent Does the Interest Burden Affect Firm Survival? Evidence from a Panel of UK Firms during the Recent Financial Crisis, *Oxford Bulletin of Economics and Statistics*, 78(4), 576-594.

Gupta, J., & Gregoriou, A. (2015). Bankruptcy and Financial Distress in US Small and Medium Sized Enterprises. *Available at SSRN 2638485*.

Hernández-Cánovas, G., & Martínez-Solano, P. (2010). Relationship lending and SME financing in the continental European bank-based system. *Small Business Economics*, 34(4), 465-482.

Hillegeist, S.A., Keating, E.K., Cram, D.P., & Lundstedt, K.G. (2004). Assessing the Probability of Bankruptcy. *Review of Accounting Studies*, 9, 5-34.

Huang, F., & He, Y. (2010). Enactment of Default Point in KMV Model on CMBC, SPDB, CMB, Huaxia Bank and SDB. *International Journal of Financial Research*, 1(1).

Hull,C. J., Nelken, I., & White, A. (2004). Merton's model, credit risk and volatility skews. *Journal of Credit Risk*, 1(1), 3-28.

Lee, W. C. (2011). Redefinition of the KMV model's optimal default point based on genetic algorithms - Evidence from Taiwan. *Expert Systems with Applications*, 38(8), 10107-10113.

Lee, N., Sameen, H., & Cowling, M. (2015). Access to finance for innovative SMEs since the financial crisis. *Research Policy*, 44(2), 370-380.

Lin, S.M., Ansell, J., & Andreeva, G. (2007). Merton Models or Credit Scoring: Modelling Default of a Small Business. *Credit Research Centre*, Management School & Economics, The University of Edinburgh.

McCarthy, E. (1999). Pricing IPOs: Science or science fiction?. *Journal of Accountancy*, 188(3), 51-58

Merton, R.C.(1974). On the Pricing of Corporate Debt: the Risk Structure of Interest Rates. *The Journal of Finance*, 29(2), 449-470.

Murro, P. (2010). Lending technologies in Italy: an example of hardening soft information? *Rivista Bancaria-Minerva Bancaria*, (4).

Office for National Statistics (2015). Business demography: 2014 Change in the number of UK businesses broken down by sector of the economy. *Statistical Bulletin*.

Petersen, M.A., & Rajan, R.G. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance*, 49(1), 3-37.

Reisz, A., & Perlich, C. (2004). A market based framework for bankruptcy prediction. *Working paper*, Baruch College, City University of New York.

Richardson, F.M., Kane, G.D., & Lobingier, P. (1998). The Impact of Recession on The Prediction of Corporate Failure. *Journal of Business Finance & Accounting*, 25(1) & (2).

Schæfer, D. (2003). The new Basel Accord and its impact on small and medium-sized companies. *German Institute for Economic Research*, Economic Bulletin 40, 209– 214.

Shumway, T. (2001). Forecasting Bankruptcy More Accurately: A Simple Hazard Model. *Journal of Business*, 74, 101-124.

Tudela, M., & Young, G. (2003). A Merton Model Approach to Assessing the Default Risk of UK Public Companies. *Bank of England.*

Uchida, H., Udell, G.F., & Yamori, N. (2008). How do Japanese banks discipline small and medium-sized borrowers? An investigation of the deployment of lending technologies. In: Choi, J., Dow, S. (Eds.), Institutional Approach to Global Corporate Governance: Business Systems and Beyond. Emerald Group Publishing Limited.

Vassalou, M., & Xing, Y. (2004). Default Risk in Equity Returns. *The Journal of Finance*, 59, 831–868.

Wang, Y. (2009). Structural Credit Risk Modelling: Merton and Beyond. Risk Management, 16, 30-33.

Zmijewski, M. (1984). Methodological Issues Related to the Estimation of Financial Distress Prediction Models, *Journal of Accounting Research*, 22, 59-82.

Table 1
Annual sample default rate for unlisted and listed SMEs (2004 to 2013)

Years	U-SMEs				L-SMEs			
	Active	%	Defaulted	%	Active	%	Defaulted	%
2004	20,694	94.24	1,265	5.76	188	94.95	10	5.05
2005	20,313	94.35	1,216	5.65	183	95.31	9	4.69
2006	19,671	93.66	1,331	6.34	180	94.74	10	5.26
2007	19,602	92.79	1,524	7.21	179	94.21	11	5.79
2008	18,486	91.44	1,731	8.56	175	93.58	12	6.42
2009	17,740	90.56	1,850	9.44	173	93.51	12	6.49
2010	19,001	93.82	1,251	6.18	182	93.81	12	6.19
2011	20,024	94.24	1,224	5.76	180	94.74	10	5.26
2012	20,425	93.48	1,424	6.52	183	94.82	10	5.18
2013	20,851	93.90	1,354	6.10	184	94.85	10	5.15
Total	**196,807**	93.28	**14,170**	6.72	**1,807**	**94.46**	**106**	5.54

Notes
This table reports the number of listed and unlisted SMEs and the frequency of default events for each calendar year and for each group. L-SMEs stands for listed SMEs; U-SMEs stands for unlisted SMEs.

Table 2
Descriptive analysis of accounting indicators for unlisted and listed SMEs (2004-2013)

Variable	U-SMEs				L-SMEs			
	Mean	**Median**	**Min**	**Max**	**Mean**	**Median**	**Min**	**Max**
RETA	-0.2534	0.0129	-0.6693	0.3217	-0.0103	0.0097	-0.2726	0.1164
STA	0.0090	0.0185	-0.5174	0.3818	0.0188	0.0175	-0.2425	0.2511
COV	84.5813	25.7211	0.0000	586.0000	99.5216	14.3114	0.0000	955.5000
LEV	1.5666	0.7111	-6.5701	15.0490	1.5072	0.6796	-2.9865	10.8652
LIQ	0.1469	0.0523	0.0000	0.6925	0.0639	0.0244	0.0000	0.3277
EBITTA	0.6155	0.4436	0.0000	2.2777	0.2479	0.2240	0.0000	0.7337
NIS	0.3245	0.3026	0.0000	0.8260	0.2277	0.2117	0.0000	0.6383
TLTA	1.9081	1.4201	0.3720	6.1781	2.5555	1.8502	0.9403	7.7632
CACL	1.6891	1.2481	0.2368	5.7092	2.0423	1.4876	2.0423	6.5727
CNS	0.1058	0.0315	0.0000	0.6509	0.0722	0.0242	0.0000	0.4184
CETL	1.5646	0.8511	-0.5000	7.6363	1.8128	1.2250	0.0757	6.9250
NCNW	0.3229	0.0728	-0.2500	2.0000	0.1351	0.0488	0.0000	0.6702
STDNW	1.6844	0.7783	-4.9795	13.7881	1.2710	0.6179	-1.3703	8.2434
CLNCL	19.0348	4.2694	0.2149	136.0000	14.6209	4.0173	0.3111	110.7822
WCTA	0.0929	0.1389	-1.0833	0.7641	0.1350	0.1418	0.0000	0.2696

Notes
This table reports the mean and median values of the accounting ratios for the listed and unlisted SME samples. All variables are winsorised 5% in each tail to eliminate the presence of outliers. *RETA* is the ratio of retained earnings to total assets; *STA* is the ratio of sales to total assets, *COV* is the ratio of *EBITDA* to interest expenses; *LEV* is the ratio of short-term debt to equity book value; *LIQ* is the ratio of cash to total assets; *EBITTA* is the ratio of earnings before interest and tax to total assets; *NIS* is the ratio of net income to sales; *TLTA* is the ratio of total liabilities to total assets; *CACL* is the ratio of current assets to current liabilities; *CNS* is the ratio of cash to net sales; *CETL* is the ratio of total capital employed to the total liabilities; *NCNW* is the ratio of net cash to net worth; *STDNW* is the ratio of short-term debt to net worth; *CLNCL* is the ratio of current liabilities to the non-current liabilities; and finally, *WCTA* is the ratio of working capital to total assets. L-SMEs stands for listed SMEs; U-SMEs stands for unlisted SMEs.

Table 3
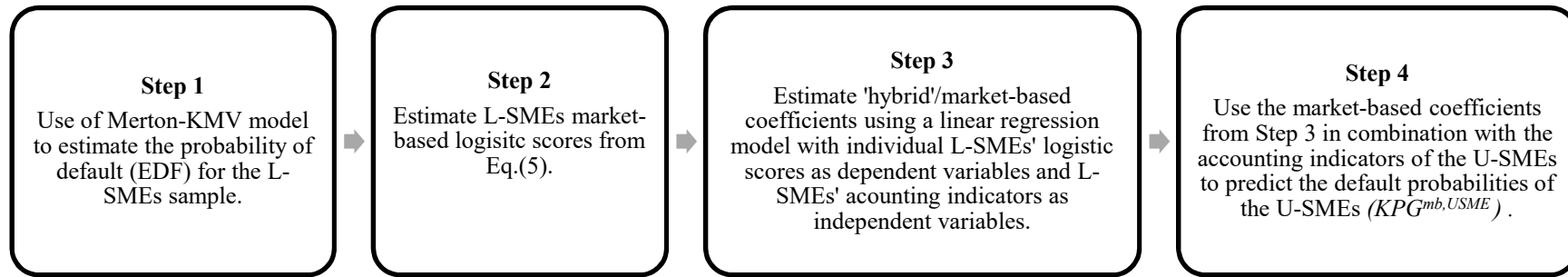Financial position of sample companies in absolute terms from 2004 to 2013 (£000s)

	U-SMEs				L-SMEs			
	Mean	**Median**	**Min**	**Max**	**Mean**	**Median**	**Min**	**Max**
Total Liabilities	4,733.8	2,876.5	0.0	15,000.0	5,437.0	3,014.8	0.0	28,400.0
Retained Earnings	1,963.8	1,567.9	-2,470.0	9,600.0	7,696.0	8,452.7	-7,100.0	13,041.2
Sales	15,352.5	11,815.8	0.0	41,000.0	35,769.1	22,815.3	1.0	41,934.6
Total Assets	5,480.4	2,956.2	0.0	17,200.0	7,755.0	6,020.4	10.0	32,900.0
Book Value of Equity	2,256.4	990.3	-1,448.0	4,100.0	3,130.1	1,934.6	-681.5	5,393.0

Notes
This table reports average, median, minimum and maximum values for key financial items for our samples of unlisted (U-SMEs) and listed SMEs (L-SMEs).
All data are extracted from Thomson's DataStream (L-SMEs) and Bureau Van Dijk's FAME (U-SMEs) covering the period 2004 to 2013.

Figure 1
'Hybrid'/Market-based default score calibration for U-SMEs

Step 1	Step 2	Step 3	Step 4
Use of Merton-KMV model to estimate the probability of default (EDF) for the L-SMEs sample.	Estimate L-SMEs market-based logisitc scores from Eq.(5).	Estimate 'hybrid'/market-based coefficients using a linear regression model with individual L-SMEs' logistic scores as dependent variables and L-SMEs' acounting indicators as independent variables.	Use the market-based coefficients from Step 3 in combination with the accounting indicators of the U-SMEs to predict the default probabilities of the U-SMEs $(KPG^{mb,USME})$.

Notes
This figure reports the process of deriving the default scores for the unlisted SMEs (U-SMEs) in our sample. L-SMEs stands for listed SMEs, *EDF* stands for Expected Default Frequency, *KPG* stands for Known Probability of Being Good, *mb* stands for market-based.

Table 4

Descriptive statistics of market-based EDFs (2004 to 2012)

	n	**Minimum**	**Maximum**	**Mean**	**Std. Err**
EDF^{LSME} (%)	1420	0.000	27.847	1.695	4.313

Notes

This table shows the summary statistics for the Expected Default Frequency (EDF) results for our L-SMEs sample for the period 2004 to 2012. L-SMEs stands for listed SMEs.

Figure 2
ROC curve using the Merton-KMV model for the L-SMEs sample (2004 – 2012)
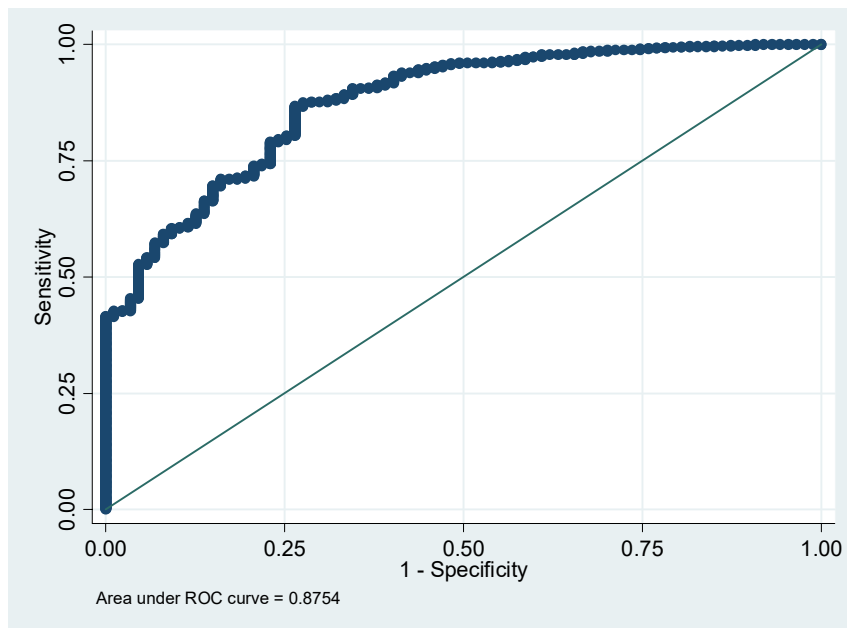
Notes
The area under the ROC curve (AUC) indicates the performance of the market-based Merton-KMV model in predicting L-SMEs default. L-SMEs stands for listed SMEs.

Table 5

Market-based (hybrid) linear regression results for the L-SMEs sample (2004-2012)

X_{it}^{LSME}	Coefficient	Robust std.err.
(n=961)		
Constant	2.1035***	0.01835
RETA	0.2387***	0.07596
STA	0.4471***	0.07540
COV	0.0006***	0.00004
LEV	-0.0024*	0.00138
LIQ	1.0852***	0.16366
EBITTA	0.1344***	0.03551
NIS	0.1494***	0.04205
TLTA	-0.1122***	0.02061
CACL	0.0275***	0.00464
CNS	0.5519***	0.11912
CETL	0.2049***	0.02220
NCNW	0.1243***	0.04352
STDNW	-0.0044***	0.00220
CLNCL	0.0002	0.00021
WCTA	0.2127***	0.07229
R-Squared	0.8568	
Adj. R-Squared	0.8545	
F-value (15, 945)	218.32***	

Notes

This table shows the results of the linear regression of market-based logistic scores for the L-SMEs sample as a dependent variable against accounting ratios as independent variables for 2004 to 2012 (Step 3 of Figure 1). All variables are winsorised 5% in each tail to eliminate the presence of outliers. X_{it}^{LSME} is the individual logistic scores for the L-SMEs at time *t; RETA* is the ratio of retained earnings to total assets; *STA* is the ratio of sales to total assets; *COV* is the ratio of EBITDA to interest expenses; *LEV* is the ratio of short-term debt to equity book value; *LIQ* is the ratio of cash to total assets; *EBITTA* is the ratio of earnings before interest and tax to total assets; *NIS* is the ratio of net income to net sales; *TLTA* is the ratio of total liabilities to total assets; *CACL* is the ratio of current assets to current liabilities; *CNS* is the ratio of cash to net sales; *CETL* is the ratio of total capital employed to the total liabilities; *NCNW* is the ratio of net cash to net worth; *STDNW* is the ratio of short-term debt to net worth; *CLNCL* is the ratio of current liabilities to the non-current liabilities; *WCTA* is the ratio of working capital to total assets. L-SMEs stands for listed SMEs. Time, industry and ownership dummies are included but not displayed. * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 6

Accounting-based logistic regression results for the U-SMEs sample (2004-2012)

$KPG_{it}^{ab,USME}$ (n=107,061)	Coefficient	Robust std.err.
Constant	1.5529***	0.03335
RETA	0.6502***	0.07466
STA	1.1106***	0.09720
COV	0.0013***	0.00009
LEV	-0.0025	0.00226
LIQ	0.8628***	0.16708
EBITTA	0.4948***	0.02936
NIS	0.2922***	0.06621
TLTA	-0.0918***	0.01653
CACL	0.1055***	0.02175
CNS	1.0165***	0.16759
CETL	0.1261***	0.01071
NCNW	0.2130***	0.02907
STDNW	-0.0271***	0.00305
CLNCL	0.0011***	0.00028
WCTA	0.3627***	0.04301
AIC	91472.13	
Log-likelihood	-45720.07	
Pseudo R-squared	0.0458	
Wald Chi-square	3783.45***	

Notes

This table reports the results of the logistic regression that uses accounting-based variables for our sample of U-SMEs during the period 2004 to 2012. All variables are winsorised 5% in each tail to eliminate the impact of outliers. The dependent variable, $KPG_{it}^{ab,USME}$, stands for *Known Probability of being Good* in line with Altman and Sabato (2007). *RETA* is the ratio of retained earnings to total assets; *STA* is the ratio of sales to total assets; *COV* is the ratio of EBITDA to interest expenses; *LEV* is the ratio of short-term debt to equity book value; *LIQ* is the ratio of cash to total assets; *EBITTA* is the ratio of earnings before interest and tax to total assets; *NIS* is the ratio of net income to net sales; *TLTA* is the ratio of total liabilities to total assets; *CACL* is the ratio of current assets to current liabilities; *CNS* is the ratio of cash to net sales; *CETL* is the ratio of total capital employed to the total liabilities; *NCNW* is the ratio of net cash to net worth; *STDNW* is the ratio of short-term debt to net worth; *CLNCL* is the ratio of current liabilities to the non-current liabilities; *WCTA* is the ratio of working capital to total assets. U-SMEs stands for unlisted SMEs; *ab* stands for accounting-based. Time, industry and ownership dummies are included but not displayed; * significant at 10%; ** significant at 5%; *** significant at 1%.

Table 7
Descriptive statistics of default scores for U-SMEs for the 2013 hold-out sample

Variable	n	Mean	Std. Dev	Min	Max
Accounting-based default scores ($KPG_{it}^{ab,USME}$)	2,431	2.5742	0.7323	-1.3214	29.3544
'Hybrid'/Market-based default scores ($KPG_{it}^{mb,USME}$)	2,431	4.5213	1.1718	0.8765	36.6541

Notes
The table shows the aggregate difference between accounting-based and the market-based default scores. These logistic scores are used to estimate the *PD*s for the 2013 hold-out sample using both models and the accuracy rates are estimated subsequently. U-SMEs stands for unlisted SMEs.
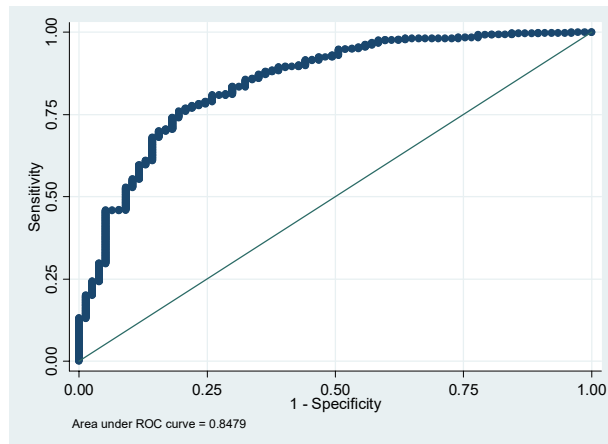
Figure 3
ROC curves for U-SMEs' hybrid- and accounting-based models using the 2013 hold-out sample and the full 2004-2012 sample period
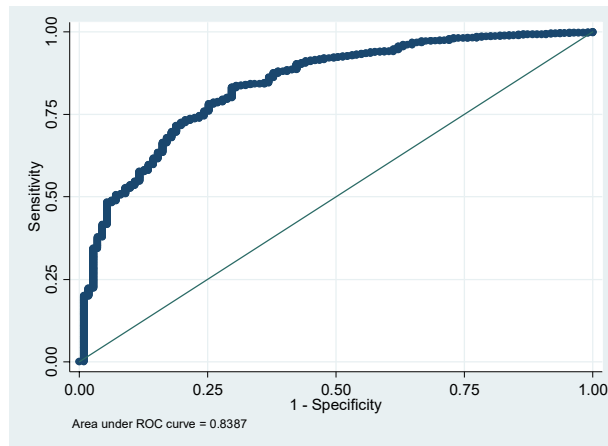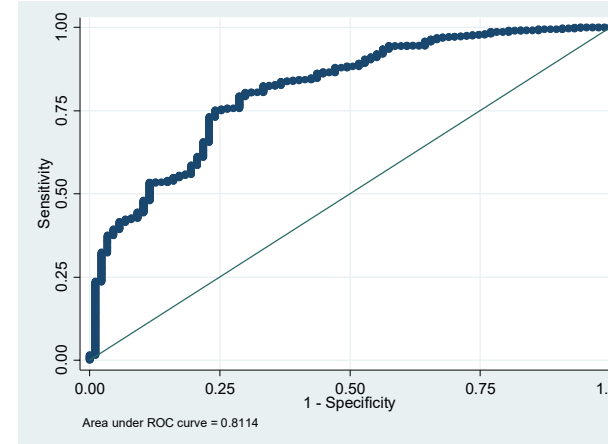

Hybrid model (2013 hold-out sample)   *Accounting-based model (2013 hold-out sample)*
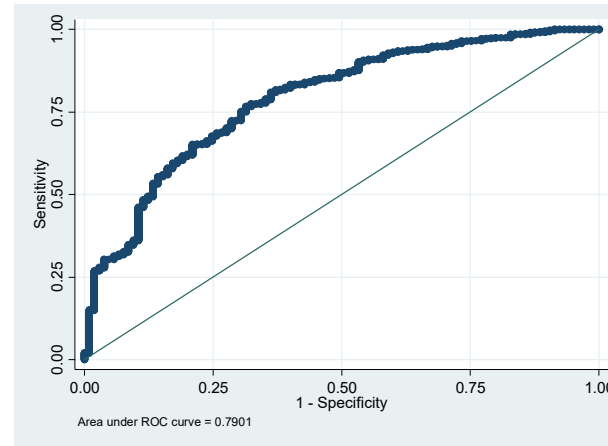
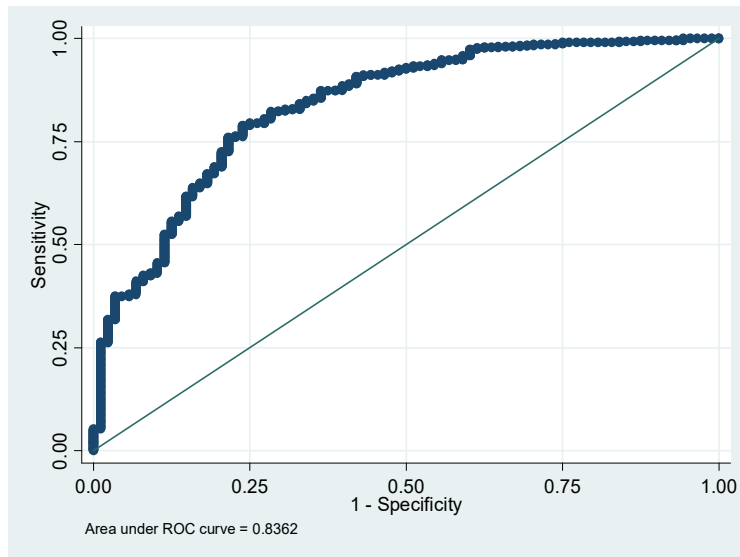*Hybrid model (Full-sample period: 2004-2012)*   *Accounting-based model (Full-sample period: 2004-2012)*
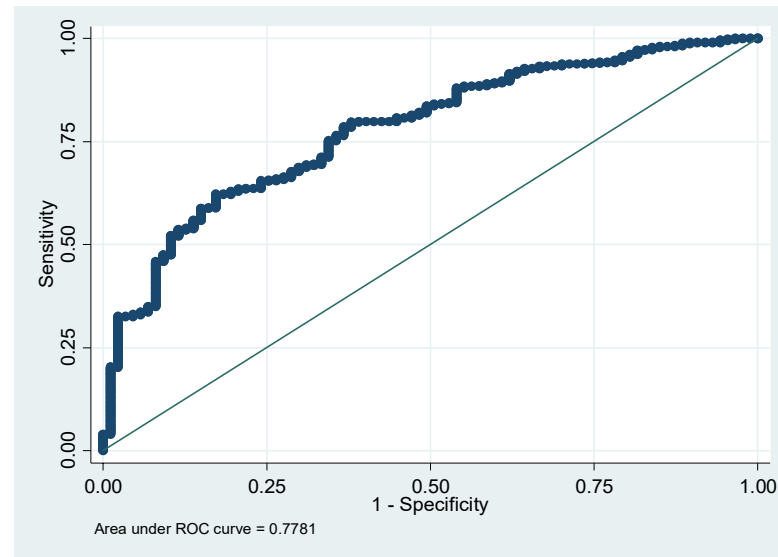
Notes
The area under the ROC curve (AUC) indicates the performance of the model in predicting U-SMEs' default. U-SMEs stands for unlisted SMEs.

Figure 4
ROC curves for U-SMEs' hybrid- and accounting-based models using the 2008 hold-out sample (Financial Crisis Period)

Area under ROC curve = 0.8362

Area under ROC curve = 0.7781

Hybrid model (2008 hold-out sample)   *Accounting-based model (2008 hold-out sample)*

Notes
The area under the ROC curve (AUC) indicates the performance of the model in predicting U-SMEs' default. U-SMEs stands for unlisted SMEs.

Table 8

Accuracy tests of the U-SMEs' hybrid and accounting-based models (2004-2012)

Year	Hybrid model	Accounting-based model
2004	85.88%	82.12%
2005	85.24%	84.21%
2006	85.35%	83.34%
2007	83.11%	80.52%
2008	82.73%	77.16%
2009	82.92%	79.33%
2010	80.09%	77.79%
2011	79.81%	80.31%
2012	84.23%	80.64%

Notes

The table shows the difference between the accuracy rates of the 'hybrid'/market-based and the accounting-based models for each calendar year in our U-SMEs' sample. Accuracy is estimated by comparing the predicted U-SMEs' defaults to the actual status of each firm in the end of each calendar year. U-SMEs stands for unlisted SMEs.