ROYAL STATISTICAL SOCIETY
DATA | EVIDENCE | DECISIONS
Journal of the Royal Statistical Society
Statistics in Society
Series A

J. R. Statist. Soc. A (2019)
182, *Part* 4, *pp.* 1205–1226

The effect of interfirm financial transactions on the credit risk of small and medium-sized enterprises

Veronica Vinciotti, Elisa Tosetti and Francesco Moscone

Brunel University London, UK

and Mark Lycett

Royal Holloway University of London, Egham, UK

[Received January 2018. Final revision July 2019]

Summary. Despite the recognized importance of interfirm financial links in determining a company's performance, only a few studies have incorporated proxies for interfirm links in credit risk models, and none of these use real financial transactions. We estimate a credit risk model for small and medium-sized enterprises, augmented with information on observed interfirm financial transactions. We exploit a novel data set on about 60000 companies based in the UK and their financial transactions over the years 2015 and 2016. We develop several network-augmented credit risk models and compare their prediction performance with that of a conventional credit risk model that includes only a set of financial ratios. We find that augmenting a default risk model with information on the transaction network makes a significant contribution to increasing the default prediction power of risk models built specifically for small and medium-sized enterprises. Our results may help bankers and credit scoring agencies to improve the credit scoring of these companies, ultimately reducing their propensity to apply excessive lending restrictions.

Keywords: Credit risk modelling; Financial transactions; Networks; Small and medium-sized enterprises

1. Introduction

Small and medium-sized enterprises (SMEs) are the engine of the economy of many countries, with over 20.7 million firms in Europe, accounting for more than 98% of all enterprises and two-thirds of all employment (Muller *et al.*, 2016). In the UK, at the beginning of 2015, SMEs accounted for 99.9% of all private sector businesses and employed 15.6 million people, representing 60% of private sector employment (White, 2015). One main feature of SMEs when compared with larger firms is that they appear to be more exposed to economic downturns and financial volatility. Such vulnerability is determined by various factors: SMEs tend to diversify less in their economic activities when compared with larger firms, they have more credit constraints for operations and cannot downsize given that they are already small (Organisation for Economic Co-operation and Development, 2009). The recent Basel Accords have encouraged international credit systems and academics to develop tools for measuring credit risk of SMEs, by allowing banks to set up an internal rating system to classify clients according to their risk. Whereas the literature has focused significantly on developing credit risk models, there is limited exploration of tools that are specifically tailored to the needs of SMEs, partly due to the lack of data from SMEs.

Address for correspondence: Veronica Vinciotti, Department of Mathematics, Brunel University London, Uxbridge, Middx, UB8 3PH, UK.
E-mail: veronica.vinciotti@brunel.ac.uk

When modelling credit risk for SMEs, one important feature to be considered is the fact that companies are not simply independent agents, but rather they exchange information, goods and services to enable production. Supply chains enable flow of materials, products and associated information downstream, while moving financial flows upstream towards manufacturers and producers of raw materials. These flows link customers and suppliers in a network of interfirm financial transactions. An unexpected event such as the reduction in the liquidity of a company or an external event may cause a propagation of shocks across the supply chain, causing connected firms to experience financial difficulties. This effect is for example shown by the empirical analysis of Soramäki *et al.* (2007), who found significant differences in the network of financial transactions between a set of commercial banks before and after the terrorist attack on September 11th, 2001. Such credit risk contagion may be amplified by the use of trade credit, which allows the purchase of products and services without immediate cash payment, potentially leading to non-payments by trade debtors (Bradley and Rubach, 2002). In a networked economy, the failure of one firm can have a snowball effect, causing failure of other companies, and in extreme cases causing an avalanche of failures, known as a bankruptcy chain (Delli Gatti *et al.*, 2006; Jacobson and von Schedvin, 2015). Hurd (2016) provided an exposition on mathematical stochastic models of contagion and cascades in the context of systemic risk and recent work proposes models that are specific to the case of credit risk (Petrone and Latora, 2018).

Despite the recognized importance of interfirm financial links in determining firms' performance, very few studies have considered incorporating data on firms' interdependence in a default prediction model. The lack of studies in this area is mostly caused by the difficulty in obtaining data on financial transactions between companies, due to privacy and confidentiality issues. For this reason, existing studies consider proxies for interfirm financial transactions, e.g. based on their geographical distance, sectorial proximity or supply chain variables (Barro and Basso, 2010; Barreto and Artes, 2016; Agca *et al.*, 2017). Network analyses, in contrast, have seen a growth in recent years, with significant progress being made both in the development of models and tools for network data and in the range of applications for which these methods have been found beneficial, from social network analyses to the modelling of financial, medical and biological networks (Kolaczyk and Csárdi, 2014). In the financial sphere, for example, social network data on customers have shown to be useful in credit scoring and fraud detection models (Van Vlasselaer *et al.*, 2015; Óskarsdóttir *et al.*, 2018).

In this paper, we exploit a unique data set from a large financial institution to study the effect of financial transactions between companies on their risk of default. Information on financial transactions enables us to observe the supplier–customer relationships between companies, thus providing us with the opportunity to study whether credit risk and financial distress are transmitted across these links. The data cover a sample of about 60000 anonymized SMEs in the UK followed over the years 2015 and 2016, and include firmographic data as well as monthly information on financial transactions between companies. We exploit this information to develop a novel credit risk model, whereby we include a number of variables that are related to the network of financial transactions as predictors of a company's credit risk, in addition to more traditional company-specific characteristics and contextual factors. In particular, for each company we consider the number and volume of transactions with its customers and suppliers in the network, the characteristics of the companies with which it has direct links and the number and volume of past transactions with companies that failed in the previous year. We estimate such a network-augmented regression model by both logistic regression and random forests. The analysis enables us to identify the characteristics of the network of transactions that are most important in determining the probability of default of companies. We then compare the prediction performance of the proposed network-augmented models with that of conventional credit risk models that include only a set of financial ratios. We find that augmenting the default risk model with information on the network of transactions makes a significant contribution to increasing the default prediction power of risk models that are built specifically for SMEs.

The remainder of the paper is structured as follows. Section 2 reviews the literature on default prediction methodologies for SMEs, analysing the most popular statistical techniques that are used to develop credit risk models for SMEs. Section 3 describes the available data. Section 4 introduces the credit risk model proposed and describes the network-based predictors that have been developed for that. Section 5 assesses its effect on the available data. Section 6 finishes with some concluding remarks, providing some suggestions for future research.

2. Background literature on credit risk models for small and medium-sized enterprises

Early work on modelling and forecasting the default of firms was primarily undertaken during the 1960s. The seminal studies in this field were those of Beaver (1966) and Altman (1968), who developed multiple-discriminant-analysis models to predict business failures by using a set of financial ratios. These works were mostly based on data referring to large corporations, as such data sets were readily available and of good quality (Altman and Narayanan, 1997). One of the first studies on credit risk for SMEs is that by Edmister (1972), who developed a multiple-discriminant-analysis model to predict small business failures, using a small sample of SMEs over the period 1954–1969 and based on 19 financial ratios. More recently, Dietsch and Petey (2004) estimated the probability of default and asset correlation in two large samples of French and German SMEs by using a one-factor ordered probit model and showed that the probability of default of SMEs is positively, rather than negatively, correlated to a firm's asset value. Finally, Glennon and Nigro (2005, 2011) examined the probability of default and time to default of loans originated under the US Small Business Administration lending programme by using survival analysis models.

An essential step in building a credit scoring model is that of selecting the predictors for inclusion in the model. Altman and Sabato (2007) developed a default prediction model using data on US firms that had annual sales of less than $65 million over the period 1994–2002. They showed that the financial ratios that are important for credit risk models on small firms are different from those of the original Altman multiple-discriminant-analysis model on large corporations. In particular, they selected five financial ratio variables, namely EBITDA (earnings before interest, taxes, depreciation and amortization) to total assets, short-term debt to total equity, retained earnings to total assets, cash to total assets and EBITDA to interest expenses. Altman *et al.* (2010) extended the Altman and Sabato (2007) model by including also qualitative information in the model, such as legal action by creditors to recover unpaid debts, company filing histories and comprehensive audit report or opinion data, which they showed improve the predictive accuracy of the model. Altman and Sabato (2007) concluded that banks would enjoy significant benefits in terms of SME business profitability if they modelled credit risk for SMEs separately from large corporations.

Although standard logistic approaches are very common in the credit scoring industry to estimate the regression coefficients of the selected predictors, Alfo' *et al.* (2005) claimed that these models often underperform when compared with other credit risk methods for SMEs. They showed how a mixed logistic model, that includes individual, firm-specific, random effects, could account for potential sources of heterogeneity among firms and was thus found to be superior to the standard logistic approaches on a rich data set of small and large Italian firms. In contrast with the statistical approaches that were mentioned above, machine learning approaches have also been considered for studying defaults of SMEs. Fantazzini and Figini (2008) proposed a random-survival forest approach and compared it with a standard logistic model in terms of forecasting performance, both in sample and out of sample. They found that, whereas random forests perform much better in terms of in-sample performance, the logistic model is preferred in terms of out-of-sample performance. Finally, Calabrese *et al.* (2016) considered penalized likelihood estimation of a credit risk model for a large sample of Italian SMEs over the period 2006–2011. Noting that employing a symmetric link function for a rare event such as default may be problematic, they considered an asymmetric generalized extreme value link function. In addition, they allowed for the functional dependence of the response on continuous covariates to be flexibly determined from the data and found that the model proposed had better forecasting properties than conventional specifications.

In the economics literature, only few recent studies have looked at the role of interfirm interaction in determining firms' default and clusters of default. Most of these references have developed theoretical models and studied under which conditions local failures can result in avalanches of shortage and failures across the network (Delli Gatti *et al.*, 2006; Weisbuch and Battiston, 2007). Some empirical studies have focused on testing the trade credit propagation hypothesis, exploring whether firms that issue more trade credit are more likely to experience a debtor failure (Jacobson *et al.*, 2013; Jacobson and von Schedvin, 2015). Few recent studies have considered incorporating information on firms' interdependence in a default prediction model. Barro and Basso (2010) proposed a model of contagion that associates the economic relationship of sectors of the economy and the geographical proximity of each pair of firms in a network of firms. Barreto and Artes (2016) built a measure of local risk of default by using ordinary kriging based on data on 9 million Brazilian SMEs observed in 2010. After including this measure as an additional explanatory variable in a logistic credit scoring model, they showed that the performance of the model improved considerably. No studies to date have used financial transactions between SMEs to capture interfirm dependences within the context of a credit risk model. This is the aim of this paper.

3. Firmographic and transaction data for small and medium-sized enterprises in the UK

We use firmographic and financial data on a sample of SMEs based in England, observed monthly between January 1st, 2015, and June 30th, 2016. SMEs in the UK are all firms that do not have more than 250 employees, a turnover smaller than £25.9 million and a total balance sheet of no more than £12.9 million. The data are provided by a large financial institution in the UK and refer to companies that hold at least one financial product (e.g. a loan or an invoice finance) with the financial institution. For each month, the data set provides information on the total amount of financial transactions between any two companies in the sample. Transaction data have been used to build a time varying network between companies, namely a directed graph, where companies are the vertices and the financial transactions represent the edges, or links. Table 1 provides a summary of the sample size, in terms of the number of companies and number of transactions, for each month available. The data set contains between 166000 and 187000 transactions for each month, with an average transaction ranging between £18000 and £26000.

Table 2 presents a summary on the percentage of companies and transactions in the sample by broad industrial group, also offering a comparison with the percentage of companies in the same sector at national level, as estimated by the Office for National Statistics. It is interesting to observe that certain industries, such as manufacturing (production) and property, are overrepresented in our data set relative to the total of the UK, whereas other sectors, such as

Table 1. Summary statistics on the number of companies and financial transactions by month

Month	*Number of companies*	*Number of transactions*	*Average amount (£)*
1	61442	171777	19327
2	59808	166426	25623
3	62966	182509	23032
4	62017	176962	20420
5	61855	175599	19148
6	63316	183275	21161
7	63678	185121	20057
8	61390	172019	18945
9	63340	182613	20553
10	63573	184025	21258
11	63268	181736	18344
12	63125	181596	22182
13	63130	175270	20429
14	62904	177611	20207
15	64545	184998	22760
16	63800	184379	20413
17	64063	184324	19796
18	64498	187028	22354

Table 2. Summary statistics on the number of companies and transactions by broad industrial grouping†

Sector	*% of companies in sample*	*% of transactions in sample (to)*	*% of transactions) in sample (from)*	*% of companies in UK*
Accommodation and food services	6.86	6.00	4.03	6.03
Business administration and support services	5.17	5.40	8.92	8.50
Construction	7.47	8.13	6.58	12.30
Agriculture, forestry and fishing	9.01	7.72	5.06	6.02
Health	11.48	10.04	6.21	4.63
Production	13.44	18.45	20.15	5.96
Property	19.44	16.89	12.64	3.71
Arts, entertainment, recreation and other services	2.51	2.46	1.50	6.86
Information and communication	1.22	1.07	0.82	8.44
Motor trade	4.40	5.96	5.87	2.99
Professional, scientific and technical	7.46	6.51	16.15	18.69
Retail	8.35	7.47	7.25	7.84
Transport and storage (including postal)	2.21	2.60	2.51	3.79
Wholesale	0.99	1.30	2.32	4.24

†The percentage of transactions is split between those that are ‘to’ and ‘from’ companies in a specific industrial group.

wholesale, are underrepresented. Such variation in the representativeness of the sample across industries may reflect differences in the need for financial products by companies belonging to different industries and with varying characteristics. In fact, relatively to other companies, firms from the manufacturing and property sector may need a wider variety of financial products that is available within a large financial institution. In terms of the number of transactions, the manufacturing and property sectors represent over a third of total outward and inward links. It is interesting to observe the presence of high spending sectors, such as the healthcare sector, *versus* high income sectors, such as business services, absorbing a large percentage of inward transactions.

The aim of this paper is to use information from this network of transactions to augment standard default prediction models, as explained further in the next section. Since some of the model predictors will need to be calculated on data from the previous year, we shall use only the companies between January 1st, 2016, and June 30th, 2016, for the empirical study. In line with other studies, we define failure as entry into liquidation, administration or receivership (Altman *et al.*, 2010). We include in the analysis all companies that were active at the beginning of the sample period, excluding companies that were dormant or became dormant during the sample period. After further removing companies which have missing values on the financial and non-financial characteristics and cleaning the data from outliers by using the function `mvBACON` in the R package `robustX` (Stahel and Maechler, 2017), the resulting data set contains 54390 companies, 1.67% of which defaulted within the 2016 sampling period considered. This is in line with the failure rate that has been reported in the literature on SMEs (Altman *et al.*, 2010).

The data set contains a subset of financial variables extracted from the accounts of firms, and provided at an annual level, as well as non-financial information related to the company and the postal district of the main trading address of each firm. For a defaulted company we consider the last set of accounts filed in the year preceding failure, whereas for non-defaulted companies we include the latest accounts. Following the literature on credit risk models for SMEs (Altman and Sabato, 2007; Altman *et al.*, 2010; Campbell *et al.*, 2008), the following variables have been considered: the age of the company, the size of the company and a set of financial ratios measuring the level of liquidity and leverage of the company. In addition, we consider the geographical location of each firm, which is defined by their postcode area and results in 25 dummy variables for the data set under consideration, and the failure rate within the broad industry groups in Table 2, which is calculated in the previous year as it requires knowledge of the class labels. The failure rate of the previous year is important to account for a possible systemic risk: when one company defaults, we generally expect negative effects on the performance of other firms in the same industry. In contrast, the failure of a firm could

Table 3. Descriptive statistics of financial and non-financial ratios for failed and not-failed companies

Variable	*Description*	*Results for failed companies*		*Results for not-failed companies*	
		Mean	*Standard deviation*	*Mean*	*Standard deviation*
Size	log(total assets)	12.428	2.555	12.965	1.803
Age	log(years)	2.620	0.660	2.606	0.736
Current assets/current liabilities	Liquidity indicator	1.425	1.676	1.741	1.963
Cash/total assets	Liquidity indicator	0.155	0.241	0.194	0.246
Networth/total liabilities	Leverage indicator	0.603	1.783	1.048	2.105
Sectoral risk	% of failed in industry sector	1.636	0.560	1.591	0.543

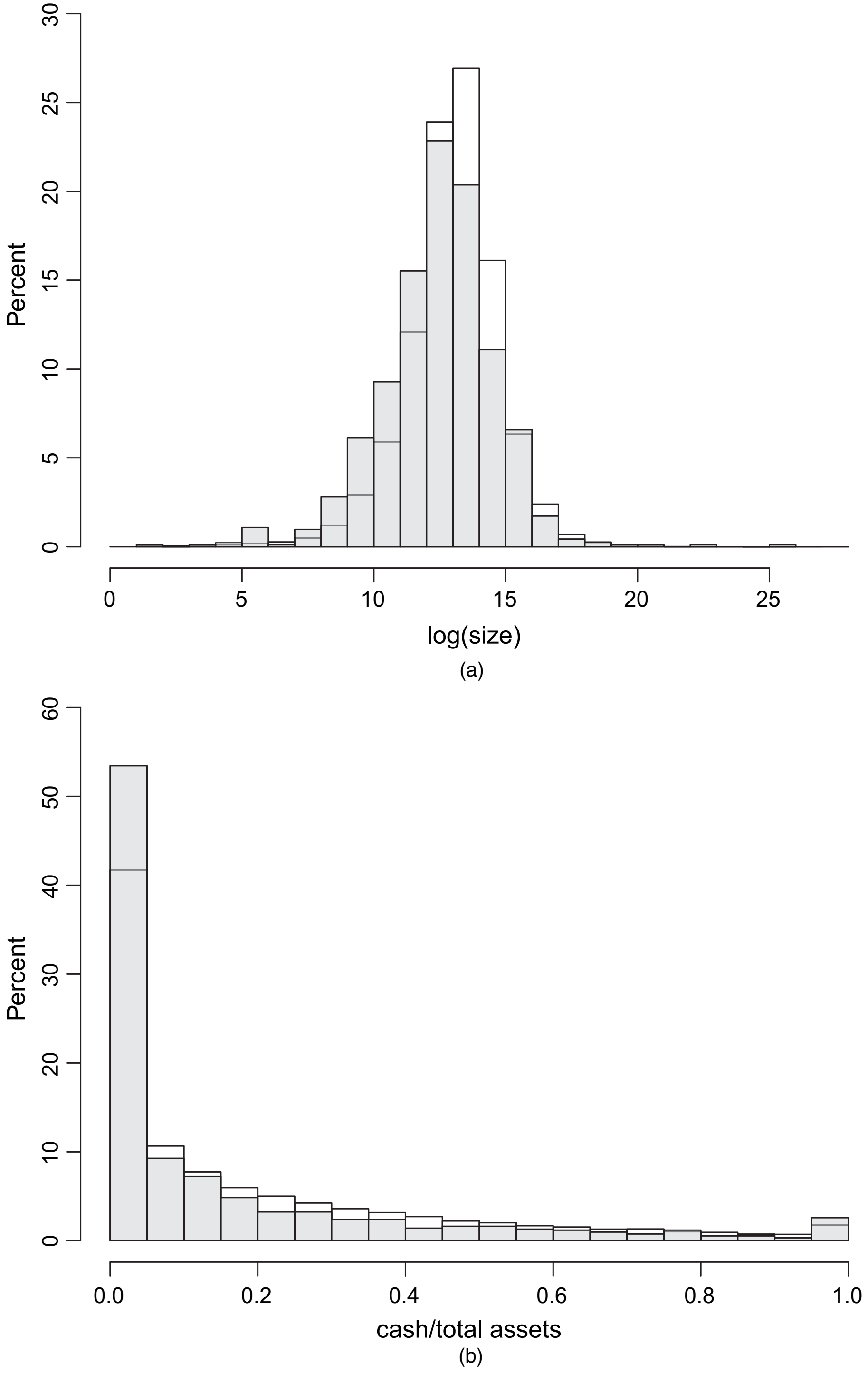


Fig. 1. Distribution of (a) size and (b) cash/total assets on the full data, split by failed (■) and not-failed (□) companies

help its competitors to gain market share. Hence, the effect of industry failure on default is likely to measure the net of these two effects in the form of large-scale (non-local) intraindustry contagion (Carling *et al.*, 2007).

Table 3 provides descriptive statistics of the financial and non-financial ratios, split by failed and not-failed companies in the final sample. As expected, firms that become insolvent tend to be smaller and they have on average better leverage indicators and worse liquidity indicators when compared with firms that are still alive. This is also shown by Fig. 1 for two representative variables. There appears to be little difference in the age of companies between the two groups, although this may be masked by a potential non-linear dependence of the probability of default on age, as shown by the empirical analysis in Section 5.

4. Network-based predictors of the credit risk of small and medium-sized enterprises

The variables in Table 3 are conventionally used in default prediction models for SMEs. Thanks to the availability of interfirm transaction data, we augment these traditional models with some novel network-based predictors, which we describe in this section. Specifically, we have included three groups of variables that capture different aspects of financial links between companies.

A first set of variables, which we denote as network characteristics, aims at defining useful summaries of the interfirm transactions. In particular, for each firm we consider the number and amount of inward and outward transactions to companies in the network as well as the amount of spending to companies outside the network, as defined by

(a) indegree, the number of companies from which transactions are received,
(b) outdegree, the number of companies to which transactions are sent,
(c) income, the amount of inward transactions,
(d) spending, the amount of outward transactions, and
(e) away-spending, the amount of spending to companies outside the network.

We expect that more numerous and larger inward transactions reduce a company's risk of default, indicating a larger pool of customers. In fact, the total income is likely to be correlated with account receivables, a variable that is not present in our data set, but that is often used as a proxy for the level of activity of the company. As for outward transactions, more spending is associated with larger investments and larger costs for row production materials, supplies and/or manufactured goods that are expected to be sold; thus it could be seen as a signal of improvement in the total activity of the company. In contrast, if for example more spending is associated with an increase in the unsold stock, then this may reduce the liquidity of the firm, thus having a negative effect on its performance. Hence, the coefficient that is attached to this variable is likely to measure the net of these two effects. Finally, total spending sent outside the transaction network gives an indication of the importance of financial transactions with companies that are external to the network and has been included in the regression models as a controller. Fig. 2 contrasts the distribution of some key network variables for failed companies against that of non-failed companies. We note that for these graphs and for the subsequent modelling all expenditure variables have been log-transformed. These graphs confirm the expectation that healthy companies are on average more connected and tend to trade more than do defaulted firms, whereas financially distressed companies tend to be poorly connected. The histogram of away-spending does not show clear differences between the failed and healthy companies, but does, however, show how a large percentage of transactions of the companies in the sample occur within the given network, with only 37% of companies having transactions with companies outside the

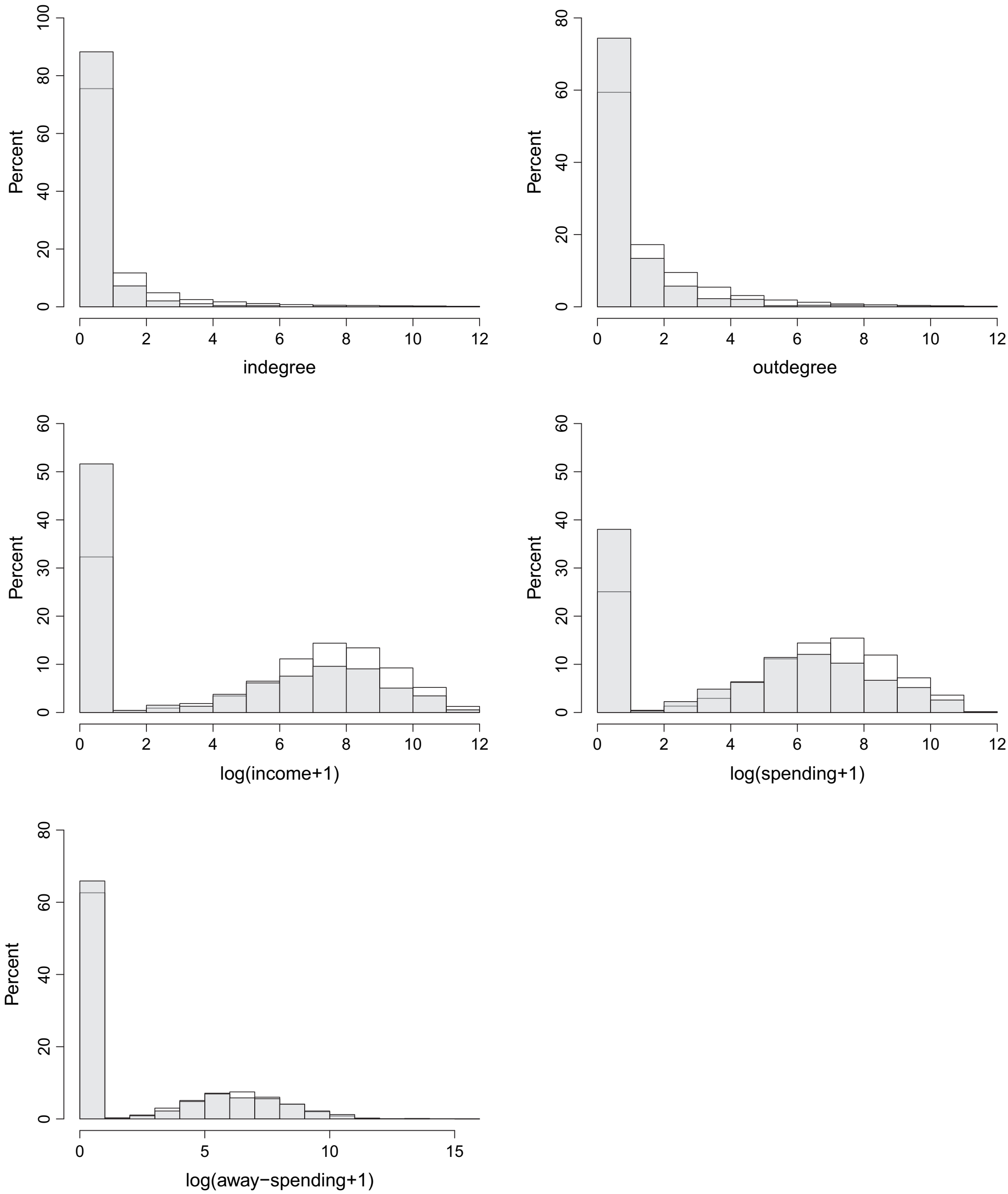


Fig. 2. Network characteristics of failed (■) *versus* not-failed (□) companies

network. Since transaction data are provided monthly whereas company characteristics and default are provided annually, we measure these variables as an average across months, where the average is calculated during the 2016 period when the company is active. A potential risk of using the latest network data is that transaction data for default companies may be present also past their time of default or may become erratic immediately before that. Fig. 3 does not show any indication of this, with most defaulted companies in the samples having data for the full

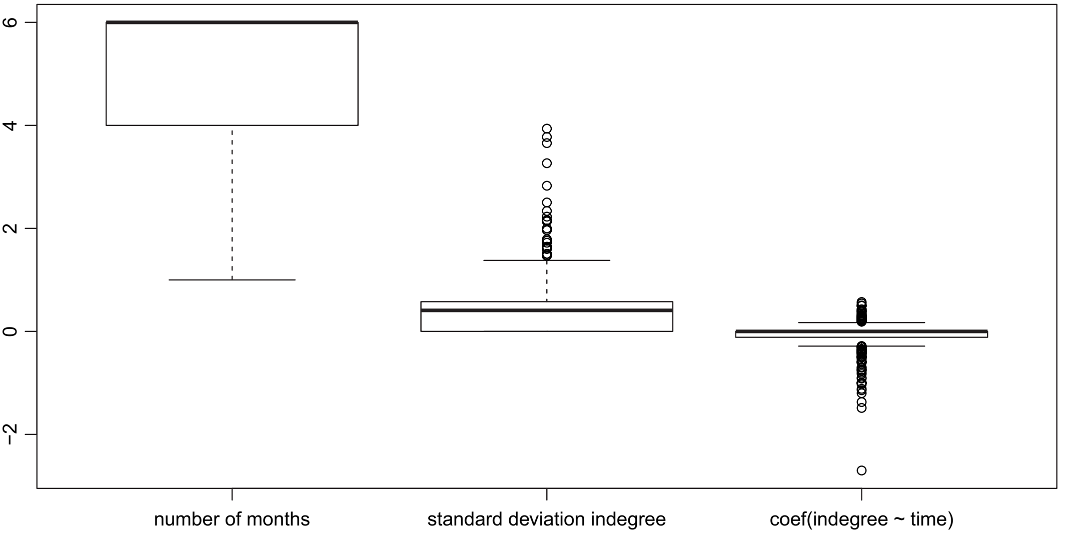


Fig. 3. For the subset of 909 defaulted companies in the sample (2016 network data), distribution of months for which network data are available, standard deviation of indegree across months and the regression coefficient of indegree over time

6-month period, showing little variability of network characteristics across these months and no indication of time-dependent variability.

As the name of indegree and outdegree implies, these measures are classical network measures in the network analysis literature (Kolaczyk and Csárdi, 2014). Similarly, income and spending are their corresponding weighted versions, i.e. indegree and outdegree calculated on the weighted adjacency matrix with the weight of an edge given by the corresponding transaction amount. The large percentage of 0s in these measures—32.4% for indegree and income, and 25.1% for outdegree and spending—correspond to the more isolated companies, that have no inward and/or outward transactions. Inspired by the network analysis literature, in the empirical section we shall also evaluate whether additional network measures, namely betweenness, closeness, eigenvector centrality and clustering coefficient (Kolaczyk and Csárdi, 2014), increase prediction performance, while also taking into consideration the disadvantages of these measures compared with the simpler indegree and outdegree in terms of computational efficiency and loss of data.

Whereas receiving income is expected to improve performance, the characteristics of firms from which a company receives income could play an important role in explaining its credit risk. Financially distressed companies are likely to have liquidity problems that make them bad payers, e.g. by not providing payment in full or by settling payments later than the due date. Thus, firms that receive income from these companies are more likely to experience financial difficulties that are linked to their customers' characteristics. To account for this effect, we include in our credit risk model a second set of network variables, which we refer to as first-order neighbourhood variables. This set is also considered by Gençay *et al.* (2015) in the context of modelling credit spreads of corporates. In the specifics of our model, we consider the average characteristics and financial ratios of Table 3, calculated on the companies that belong to the first-order inward neighbourhood of a given company. For example, we expect that trading with younger and smaller businesses facing liquidity problems has a negative effect on a company's performance. As before, also these variables are computed as averages across months, as the network may change across the various months. In Section 5, we shall explore with both simple

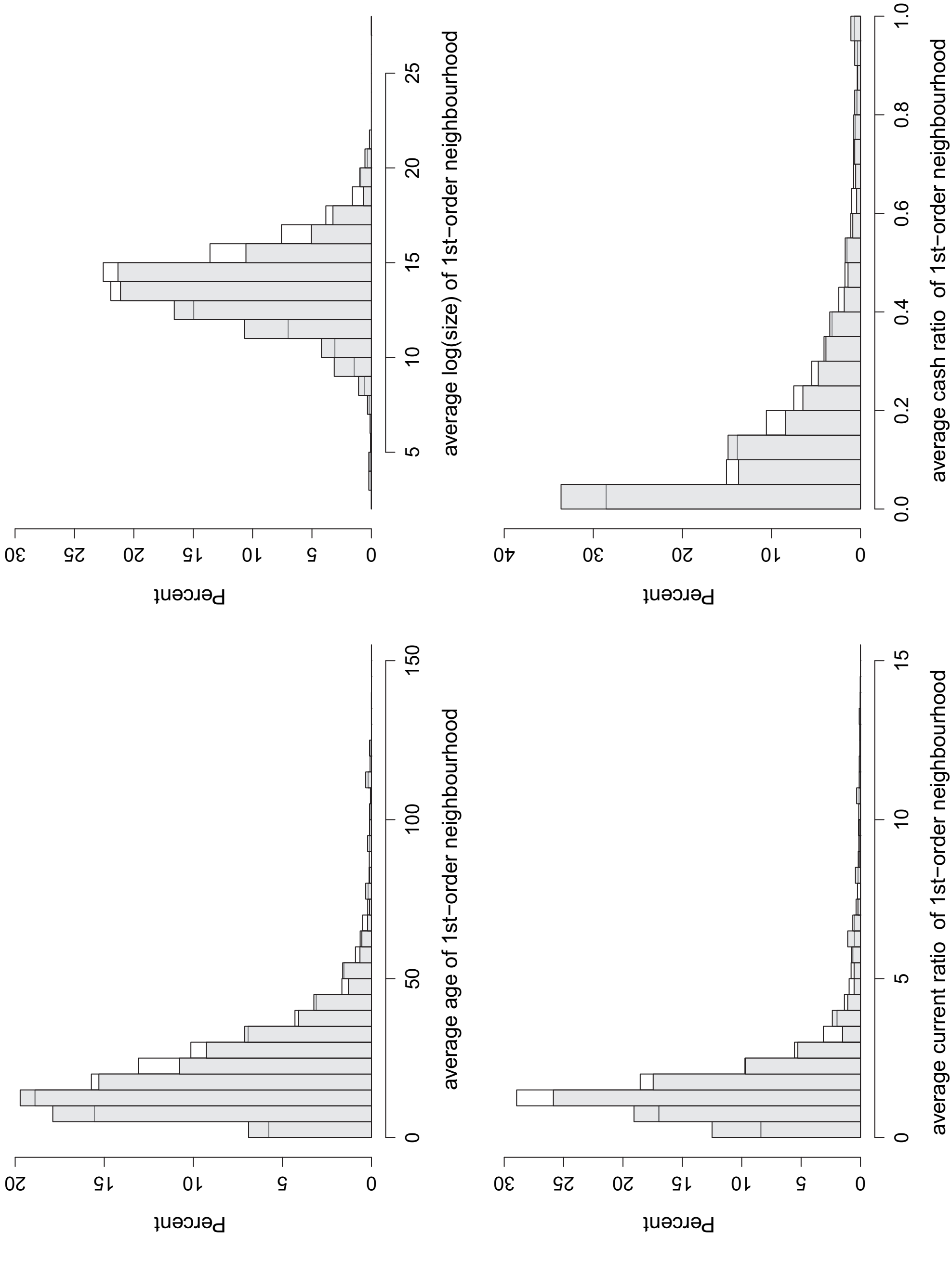


Fig. 4. First-order neighbourhood characteristics of failed (■) *versus* not-failed (□) companies

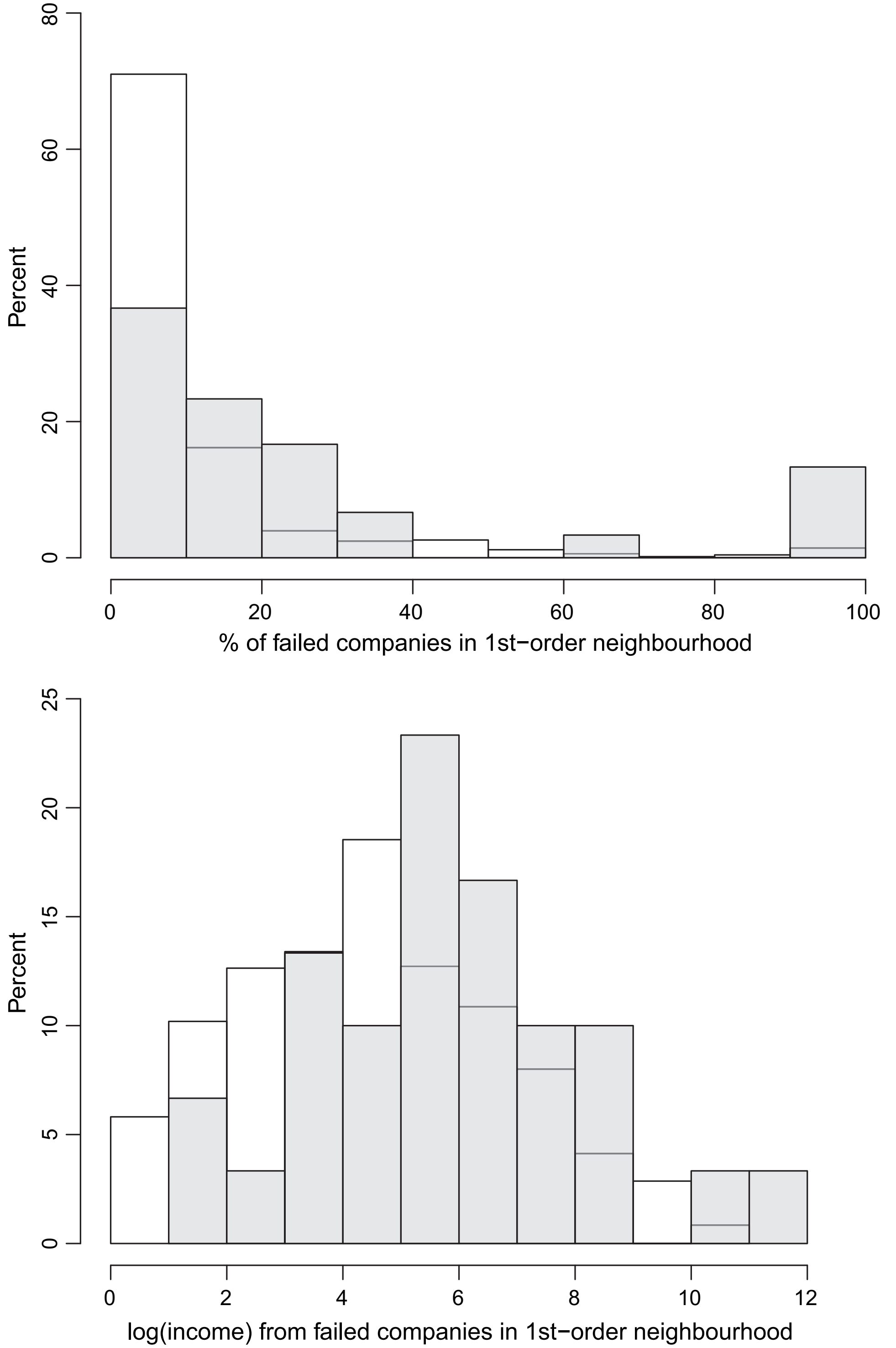


Fig. 5. Risk-based first-order neighbourhood characteristics of failed (■) *versus* not-failed companies (□)

averages and time-weighted averages, where more recent data will receive a bigger weight. For companies that have zero indegree, these variables are missing and have been imputed once by using the predictive mean matching imputation method in the R package `mice` (van Buuren and Groothuis-Oudshoorn, 2011). Fig. 4 contrasts the distribution of some key first-order neighbourhood variables for failed companies against that of non-failed companies. From Fig. 4, there is an indication that defaulted companies tend to trade with companies that have a higher risk of default themselves, being younger and smaller, and with worse liquidity and leverage indicators. In Section 5, we shall experiment also going beyond a first-order neighbourhood, by evaluating the effect of second-order neighbourhood variables on the credit risk of a company. However, in this case, the number of missing values will increase further, since companies for whom any first-order neighbours have indegree 0 will now also produce missing values.

Using the same rationale as for the second set of variables, and inspired by the literature on weighted relational neighbour classifiers (Macskassy and Provost, 2007), we include a third set of variables that are more directly linked to companies that belong to the first-order inward neighbourhood and that have failed. To avoid using the current class labels in the construction of these predictors, these measures are calculated on network data from the previous year. Specifically, we consider the inclusion of the following risk-based first-order neighbourhood variables:

(a) %failed, the percentage of failed companies in the first-order neighbourhood,
(b) %failedweighted, the percentage of failed companies in the first-order neighbourhood, weighted by the transaction value, and
(c) incomefailed, the average income from failed companies in the first-order neighbourhood.

The credit risk of a company could be influenced by transactions with default companies in two ways. First, it could be negatively affected by the financial situation of customer companies in the year before failure, given that these are likely to be bad payers. In addition, the credit risk of a company could deteriorate because of the loss of customers who default. The percentage and average amount of past (inward) transactions with defaulted companies represent the channel of contagion in failure of firms across the network. Understanding whether a firm has been engaging with healthy or unhealthy firms may prove to be crucial in better predicting probabilities of default. By focusing on transactions with companies that defaulted in the previous year, we allow for a temporal lag in the effect of defaulted customers on the financial health of a company. Fig. 5 shows how defaulted companies have a significantly higher percentage of links with defaulted companies, as well as a higher total volume of inward transactions.

In the next sections, we use these chosen variables within a credit risk model and assess their effect on the prediction of default. We do so both by using a standard logistic regression approach, which transparently quantifies the effect of each individual predictor on credit risk and which is traditionally used in the credit risk industry, and a random-forest approach (Breiman, 2001), which is a more flexible machine learning method based on an ensemble of decision trees which has been found to perform well in many classification tasks and under various benchmark studies (Couronné *et al.*, 2018).

5. Assessing the effect of company and network data on credit risk

Table 4 shows the results of a logistic regression analysis. We contrast the results of a default prediction model with no information on the network of transactions (model I) with those of the model augmented with network characteristics (model II), first-order neighbourhood characteristics (model III) and risk-based first-order neighbourhood characteristics (model IV).

Table 4. Four credit risk logistic models fitted on the UK sample of SMEs, with incremental sets of predictors†

Variable	*Results for model I*		*Results for model II*		*Results for model III*		*Results for model IV*	
	Parameter	*Standard deviation*	*Parameter*	*Standard deviation*	*Parameter*	*Standard deviation*	*Parameter*	*Standard deviation*
Company characteristics								
Size	−0.196‡	0.018	−0.150‡	0.019	−0.138‡	0.021	−0.138‡	0.021
Age	1.647‡	0.307	1.671‡	0.307	1.669‡	0.307	1.658‡	0.307
Age^2	−0.264‡	0.056	−0.258‡	0.056	−0.258‡	0.057	−0.255‡	0.057
Current assets/liabilities	0.039	0.030	0.045	0.029	0.046	0.029	0.047	0.029
Cash/total assets	−0.971‡	0.163	−0.885‡	0.162	−0.865‡	0.167	−0.852‡	0.167
Networth/total liabilities	−0.120‡	0.032	−0.123‡	0.031	−0.124‡	0.031	−0.125‡	0.031
Sectorial risk	0.095	0.060	−0.026	0.062	−0.032	0.062	−0.032	0.062
Network characteristics								
Indegree			−0.179‡	0.044	−0.177‡	0.044	−0.177‡	0.044
Outdegree			−0.066	0.030	−0.066	0.030	−0.067	0.031
Income			−0.060‡	0.011	−0.061‡	0.011	−0.060‡	0.011
Spending			−0.054‡	0.012	−0.054‡	0.012	−0.053‡	0.012
Away-spending			0.037‡	0.012	0.038‡	0.012	0.037‡	0.012
First-order neighbourhood characteristics								
Size					−0.036	0.019	−0.036	0.019
Age					0.000	0.002	0.000	0.003
Current assets/liabilities					−0.028	0.025	−0.030	0.025
Cash/total assets					0.056	0.180	0.053	0.180
Networth/total liabilities					0.021	0.022	0.022	0.022
Risk-based first-order neighbourhood characteristics								
% failed							2.124‡	0.469
Akaike information criterion AIC	9086.2		8894		8897.7		8885.4	

†Model I, standard default model which does not use information on financial transactions; model II, additional information on the overall topology of the network; model III, additional company characteristics of the first-order neighbourhood of a given company; model IV, additional information on the credit risk of companies in the first-order neighbourhood.
‡Significance at the 1% level.

Focusing on the upper panel of Table 4, the coefficients that are attached to 'Size' across all regressions indicate that, as expected, larger companies have lower probabilities of default. The coefficients that are associated with 'Age' suggest a clear quadratic dependence of default risk on age, with an increase in the probability of default from the initial 'honeymoon' period (less than 4 years) to the later period and a subsequent decrease as the company becomes more established. 'Cash/total assets' and 'Networth/total liabilities' have a negative and significant effect on the dependent variable showing that companies with higher cash reserves relative to total assets as well as a lower level of short- and long-term indebtedness with respect to the capital owned are less likely to default. Across all regressions, the current asset ratio is not statistically significant, suggesting that the ability to pay short-term and long-term obligations with its current assets is not an important determinant of a firm's default once all the other variables have been accounted for. The same is observed for the 25 variables indicating the geographical location of the company, which are therefore not reported in Table 4. The results reported in columns II–IV, which use information on the network of transactions, show, across all regressions, a negative and statistically significant effect on the probability of default for both the number of inward links (indegree) and average income and spending: the higher the number and volume of inward and outward transactions, the lower the probability for the firm to default. These findings indicate that, overall, transacting more with other companies is beneficial for a company's financial health, pointing at the network variables as proxies for the volume of activity of a company. Focusing on the characteristics of the first-order (inward) neighbourhood (columns III and IV), the results show no statistically significance effect of these on the probability of default once the company and network characteristics have been accounted for. Finally, the results from model IV, which includes the percentage of financial transactions with companies that failed in the previous year, indicate that the higher the previous year's interaction with failed companies the higher the current probability of default, thus supporting the hypothesis of contagion effects in default across the transaction network. The same conclusion was reached by using either the variable incomefailed or %failedweighted in place of %failed, because of the high correlation between these variables (0.70 between %failed and incomefailed and 0.96 between %failed and %failedweighted).

The results are on the whole supported also by the analysis using random forests in place of logistic regression. Compared with logistic regression, random forests requires tuning for some of the parameters. We used the function `OOBCurve` in the homonymous R package (Probst and Boulesteix, 2018) to tune the number of trees, whereas we use the latest implementations in the `tuneRanger` R package (Probst *et al.*, 2019) to tune the number of variables randomly sampled as candidates at each split, the minimal size of terminal nodes and the fraction of observations to sample (function `tuneRanger`). In all cases, we choose the area under the curve (AUC) as the performance criterion for tuning. For the more complex model, the optimal parameters are given by 1000 trees, 17 variables sampled as candidates at each split, a minimum sample size of 666 for the terminal nodes and a fraction of 61.09% of observations to sample. Fig. 6(a) shows the ranking of the variables in terms of Gini impurity from the fitted random-forest model with all variables included. Although the results confirm those from the other analyses for variables such as 'Size' and 'Cash/total assets', a main difference to note is on the importance that is placed now on the characteristics of companies in the first-order neighbourhood of a given company. In particular, it is interesting to observe that size and liquidity (as measured by the current ratio) of neighbours decrease the risk of default. Hence, according to this analysis, being the supplier of companies that are large and well established, with good cash reserves, is quite important for the financial health of a company, perhaps guaranteeing a more stable flow of money, and an upward shift in the total activity of the company. Finally, Fig. 6(b) confirms a clear quadratic

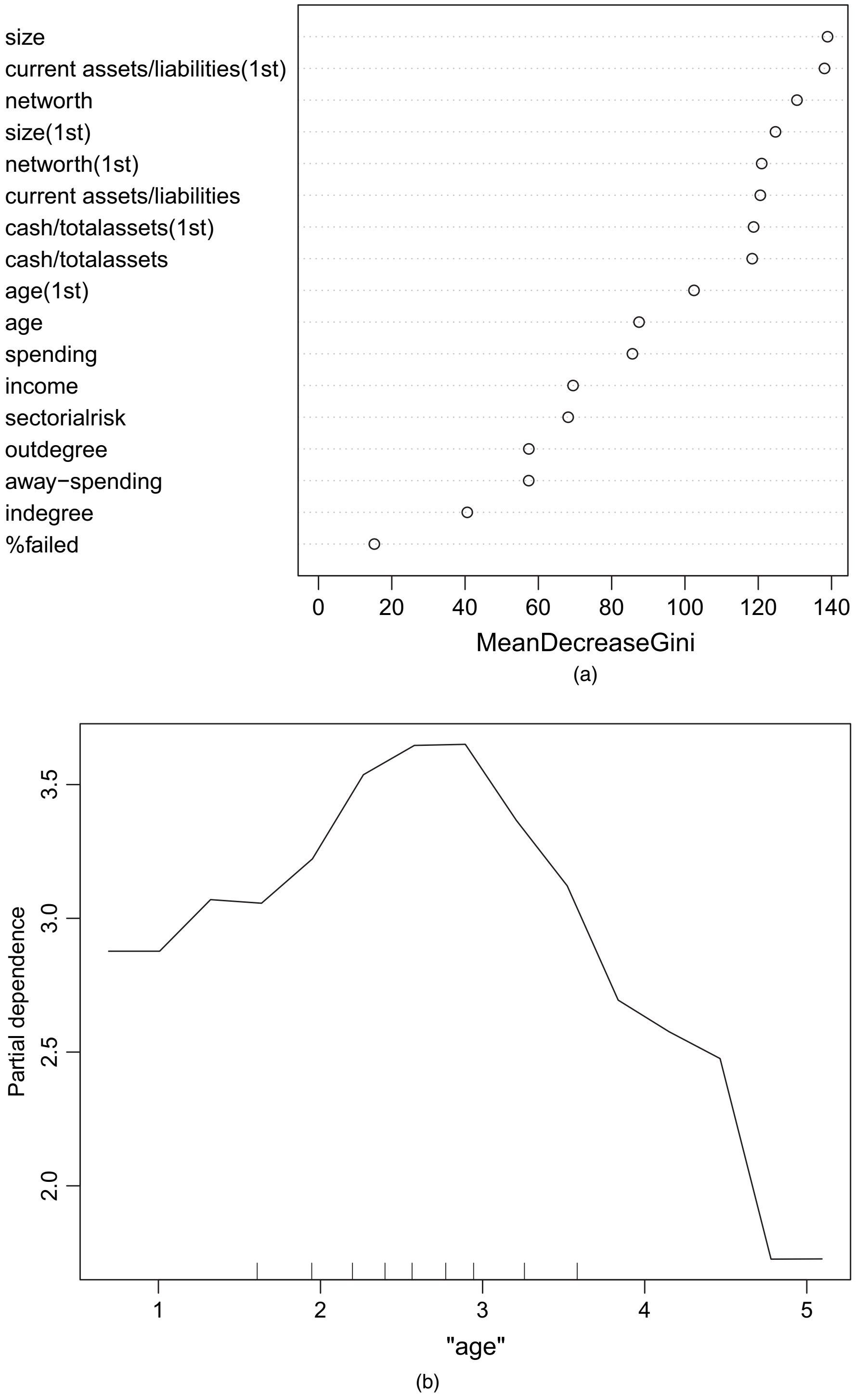
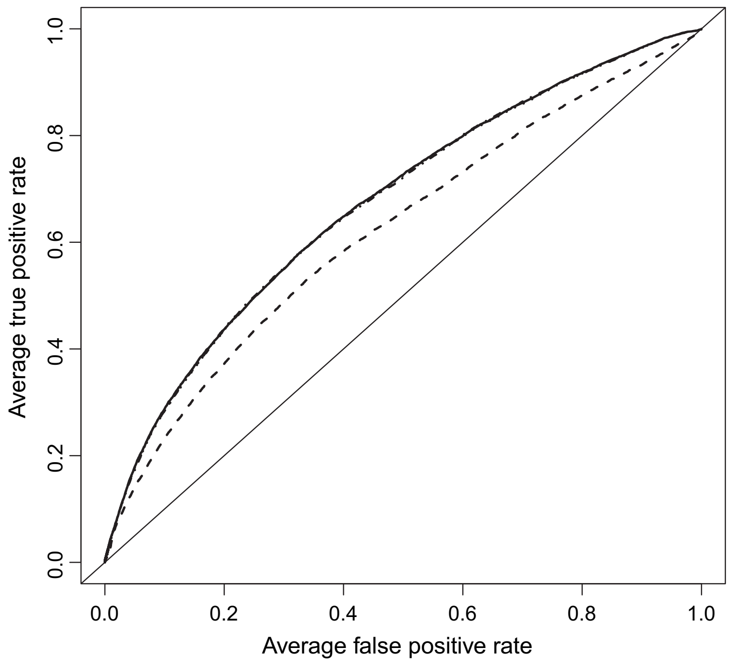


Fig. 6. Random forests for model IV fitted on the full data: (a) variable importance based on Gini impurity; (b) partial dependence plot on age

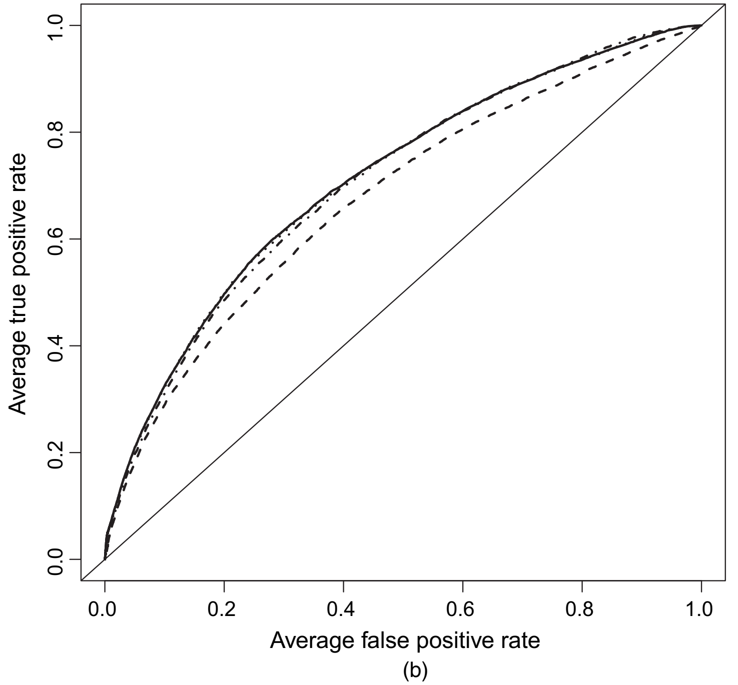
dependence of failure on age, which can be picked up by a non-linear classification model like random forests without the need for additional predictors.

To evaluate the forecast accuracy of the network-augmented models, we repeatedly split the data into a training and test sample, with 75% of the data in the training sample and the remaining 25% used for testing. Fig. 7 reports the average receiver operating characteristic curves across 50 splits for the logistic regression and random-forest approaches. For random forests, we set the number of trees to 1000 and tune the parameters on the training data at each split. Table 5 further summarizes the performance in terms of the AUC and two additional performance measures: the Kolmogorov–Smirnov statistic and the H-measure (Hand, 2009). The measures are summarized across the 50 training–test splits and are presented both for the in-sample and the out-sample cases. Out-of-bag in-sample predictions are used for the calculation of the in-sample measures from the fitted random forest. The results for the standard credit risk model show a relatively low performance (median AUC 0.62 for logistic regression and 0.67 for random forests), which is, however, in accordance with the literature on credit risk models for SMEs; for example an AUC of 0.67 was reported by Altman *et al.* (2010), whose model does, however, include some additional financial ratios that are known to be important determinants of credit risk, such as those related to the ability of the firm to generate profits. The poor performance for logistic regression could also be due to the skewness and bimodal distribution of some of the predictors, in most cases caused by the large percentage of companies with zero indegree or outdegree; for example see Fig. 2. Combined with the low failure rate which is typical of these data (1.7%), robustness checks did not show an improvement in performance (e.g. a standard logistic regression on all data removing 32.4% of companies with zero indegree, which includes also 22.2% of companies with zero away-spending and 10% of companies with zero outdegree, returned an AUC of 0.59). Although the performance of the models could be improved with a larger data set and additional financial ratios, the results give a clear indication that information on the transaction network improves the overall performance of the models considered; for example there is a significant increase in the AUC from 0.62 to 0.67 for logistic regression on the out-of-sample data (DeLong p-value 4.42×10^{-5} (Sun and Xu, 2014)) and from 0.67 to 0.72 for random forests (DeLong p-value 0.016). In addition, both analyses show how the network characteristics contribute to the biggest improvement in prediction accuracy and that there is little gain in adding the first-order neighbourhood variables (the DeLong p-values in Table 5 are not significant when comparing models II–IV). As a further benchmark, we consider a simple segmentation approach, based on the most significant variable detected both by logistic regression and random forests. Using the average split for the variable Size from the random-forest approach (12.55), we split the data set into small and large companies and run the analysis on the two subsets separately. The analysis reveals no difference in performance of the logistic regression model on the larger companies *versus* the smaller companies (the AUC for model IV was equal to 0.65 on small companies *versus* 0.64 on large companies), although the separate analyses show a slightly lower performance than the analysis on the full data set (AUC equal to 0.67).

Further analyses were conducted by considering additional or alternative predictors. As a first analysis, we extended the model by adding second-order neighbourhood variables. This required missing data imputation for a larger number of observations, since companies for whom any first-order neighbours have indegree 0 will now also produce missing values (approximately 12%), in addition to all companies whose first-order neighbourhood variables are missing (approximately 32%). Possibly also as a result of missing data in these variables, no improvement was observed in the model once these variables had been included (AUC equal to 0.67 for logistic regression and 0.71 for random forests). As a second analysis, no improvement was observed when considering time-weighted versions of the first-order neighbourhood variables, with more recent months



(a)



(b)

Fig. 7. Receiver operating characteristic curves for default prediction models under the inclusion of various sets of variables, calculated on test data and averaged across 50 75%–25% training–test splits (- - - -, company characteristics; · – · – · –, plus network characteristics; · · · · · · · ·, plus characteristics of neighbours; ———, plus failure of neighbours): (a) logistic regression; (b) random forests

Table 5. In- and out-of-sample performance measures for the four models in Table 4, reported as the median and 95% empirical confidence intervals across 50 75%–25% training–test splits

	Results for model I		*Results for model II*		*Results for model III*		*Results for model IV*	
	In sample	*Out of sample*	*In sample*	*Out of sample*	*In sample*	*Out of sample*	*In sample*	*Out of sample*
Logistic regression								
Kolmogorov–Smirnov								
Median	0.23	0.21	0.29	0.27	0.29	0.26	0.29	0.26
95% confidence interval	(0.21,0.25)	(0.17,0.25)	(0.27,0.31)	(0.23,0.31)	(0.28,0.31)	(0.23,0.31)	(0.28,0.31)	(0.23,0.31)
H-measure								
Median	0.08	0.07	0.13	0.11	0.13	0.11	0.13	0.12
95% confidence interval	(0.07,0.09)	(0.05,0.09)	(0.12,0.14)	(0.08,0.14)	(0.12,0.15)	(0.08,0.14)	(0.12,0.15)	(0.08,0.15)
AUC								
Median	0.65	0.62	0.69	0.67	0.69	0.67	0.69	0.67
95% confidence interval	(0.63,0.66)	(0.59,0.65)	(0.68,0.70)	(0.65,0.70)	(0.69,0.70)	(0.65,0.70)	(0.69,0.70)	(0.65,0.70)
DeLong								
Model I			7.82×10^{-5}		7.20×10^{-5}		4.42×10^{-5}	
Model II					0.45		0.30	
Model III							0.18	
Random forests								
Kolmogorov–Smirnov								
Median	0.27	0.27	0.32	0.33	0.32	0.33	0.33	0.33
95% confidence interval	(0.25,0.30)	(0.24,0.34)	(0.30,0.34)	(0.27,0.37)	(0.30,0.35)	(0.28,0.39)	(0.31,0.35)	(0.28,0.38)
H-measure								
Median	0.11	0.12	0.14	0.15	0.15	0.16	0.16	0.15
95% confidence interval	(0.10,0.13)	(0.08,0.16)	(0.13,0.16)	(0.11,0.18)	(0.14,0.17)	(0.13,0.20)	(0.14,0.17)	(0.13,0.20)
AUC								
Median	0.68	0.67	0.71	0.71	0.71	0.71	0.72	0.71
95% confidence interval	(0.67,0.69)	(0.64,0.71)	(0.70,0.72)	(0.68,0.73)	(0.70,0.72)	(0.69,0.74)	(0.71,0.72)	(0.69,0.74)
DeLong								
Model I			0.018		0.017		0.016	
Model II					0.335		0.338	
Model III							0.481	

having a linearly larger weight than past months (AUC equal to 0.68 for logistic regression and 0.72 for random forests). Finally, we considered the inclusion of additional network measures, namely betweenness, closeness, eigenvalue centrality and clustering coefficient (Kolaczyk and Csárdi, 2014), all calculated in their normalized form by using the `igraph` R package. Notwithstanding the computational issues in calculating betweenness and closeness for big networks (we use the function `estimate_betweenness` and `estimate_closeness` with a cut-off of 5 on the maximum path length), possible identifiability problems in calculating eigenvector centrality from asymmetric adjacency matrices (which were not observed in our analysis using the function `eigen`), and finally additional missing values in the calculation of the clustering coefficient, which is not defined for all nodes with zero or one neighbours, we found no or only moderate improvement in the use of these additional measures (AUC equal to 0.67 and 0.73 for logistic regression and random forests respectively, when betweenness, closeness and eigenvalue centrality were added, and AUC equal to 0.62 and 0.67 for logistic regression and random forests respectively, when transitivity was further included and 40% of missing data removed).

6. Concluding remarks

Interfirm financial transactions capture the strength of supplier–customer relationships, providing one possible channel of contagion of financial shocks across companies. Despite their importance, researchers struggle to gather information on interfirm financial links, because of privacy and confidentiality issues. To our knowledge, this paper is a first attempt to show the effect of intercompany financial transactions on the default of SMEs.

In line with existing literature, our results show that firm's specific characteristics, such as their size and financial indicators, play an important role in understanding the probability for a company to survive. The novel contribution of this paper is to identify a set of important network characteristics that improve the prediction power of credit risk models. In particular, we find that the total number and volume of inward and outward transactions of a company are key determinants of their probability of surviving. In making these statements, we accept the limitation that the data that are used in this paper cover only two years, and that they provide only a selection of company financial ratios. Future work will on the one hand extend the analysis to time series data to investigate the temporality of contagion better, and on the other hand match the available data with further information on company accounts, to carry out a more accurate forecasting analysis.

Given the low failure rate that is typical of these data, more extensive analyses on larger data sets are needed to validate the improvement that network data can bring to credit risk modelling. The implications of this are numerous. By incorporating information on the network in which SMEs are embedded in a credit risk model, bankers and credit scoring companies may better capture the complexity of failure processes and thus improve their rating models of SMEs. Such information could also help SMEs to select their network better and to reduce the risk of engaging in business with other, non-healthy, companies. Finally, central bankers can use the network effect on the credit risk of firms to calibrate their macroeconomic models for monetary policy purposes better.

Acknowledgements

The authors acknowledge financial support from the Engineering and Physical Sciences Research Council (grant EP/L021250/1). We thank the financial organization that provided the data and Dr George Foy for assisting with the data retrieval.

References

Agca, S., Babich, V., Birge, J. R. and Wu, J. (2017) Credit risk propagation along supply chains: evidence from the CDS market. *Research Paper 3078752*. Georgetown McDonough School of Business, Washington. (Available from `https://ssrn.com/abstract=3078752`.)

Alfo', M., Caiazza, S. and Trovato, G. (2005) Extending a logistic approach to risk modeling through semiparametric mixing. *J. Finan. Serv. Res.*, **28**, 163–176.

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

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

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

Altman, E. I., Sabato, G. and Wilson, N. (2010) The value of non-financial information in small and medium-sized enterprise risk management. *J. Credt Risk*, **6**, 1–33.

Barreto, G. and Artes, R. (2016) Spatial dependence in credit risk and its improvement in credit scoring. *Eur. J. Oper. Res.*, **249**, 517–524.

Barro, D. and Basso, A. (2010) Credit contagion in a network of firms with spatial interaction. *Eur. J. Oper. Res.*, **205**, 459–468.

Beaver, W. H. (1966) Financial ratios as predictors of failure. *J. Accountng Res.*, **4**, 71–111.

Bradley, D. B. and Rubach, M. J. (2002) Trade credit and small business: a cause of business failures. *Technical Report*. Small Business Advancement National Center, University of Central Arkansas, Conway.

Breiman, L. (2001) Random forests. *Mach. Learn.*, **45**, 5–32.

van Buuren, S. and Groothuis-Oudshoorn, K. (2011) mice: multivariate imputation by chained equations in R. *J. Statist. Softwr.*, **45**, 1–67.

Calabrese, R., Osmetti, S. and Marra, G. (2016) Bankruptcy prediction of small and medium enterprises using a flexible generalized extreme value model. *J. Oper. Res. Soc.*, **67**, 604–615.

Campbell, J. Y., Hilscher, J. and Szilagyi, J. (2008) In search of distress risk. *J. Finan.*, **63**, 2899–2939.

Carling, K., Jacobson, T., Linde, J. and Roszbach, K. (2007) Corporate credit risk modelling and the macroeconomy. *J. Bnkng Finan.*, **31**, 845–868.

Couronné, R., Probst, P. and Boulesteix, A. (2018) Random forest versus logistic regression: a large-scale benchmark experiment. *BMC Bioinform.*, **19**, article 270.

Delli Gatti, D., Gallegati, M., Greenwald, B., Russo, A. and Stiglitz, J. E. (2006) Business fluctuations in a credit-network economy. *Physica* A, **370**, 68–74.

Dietsch, M. and Petey, J. (2004) Should SME exposures be treated as retail or corporate exposures?: a comparative analysis of default probabilities and asset correlations in French and German SMEs. *J. Bnkng Finan.*, **28**, 773–788.

Edmister, R. (1972) An empirical test of financial ratio analysis for small business failure prediction. *J. Finan. Quant. Anal.*, **7**, 1477–1493.

Fantazzini, D. and Figini, S. (2008) Random survival forest models for SME credit risk measurement. *Methodol. Comput. Appl. Probab.*, **11**, 29–45.

Gençay, R., Signori, D., Xue, Y., Yu, X. and Zhang, K. (2015) Economic links and credit spreads. *J. Bnkng Finan.*, **55**, 157–169.

Glennon, D. and Nigro, P. (2011) Evaluating the performance of static versus dynamic models of credit default: evidence from long-term small business administration-guaranteed loans. *J. Credt Risk*, **7**, 3–35.

Glennon, D. C. and Nigro, P. (2005) Measuring the default risk of small business loans: a survival analysis approach. *J. Money Credt Bnkng*, **37**, 923–947.

Hand, D. J. (2009) Measuring classifier performance: a coherent alternative to the area under the ROC curve. *Mach. Learn.*, **77**, 103–123.

Hurd, T. (2016) *Contagion!: Systemic Risk in Financial Networks*. Basel: Springer Nature.

Jacobson, T., Lind, J. and Roszbach, K. (2013) Firm default and aggregate fluctuations. *J. Eur. Econ. Ass.*, **11**, 945–972.

Jacobson, T. and von Schedvin, E. (2015) Trade credit and the propagation of corporate failure: an empirical analysis. *Econometrica*, **83**, 1315–1371.

Kolaczyk, E. and Csárdi, G. (2014) *Statistical Analysis of Network Data with R*. New York: Springer.

Macskassy, S. A. and Provost, F. (2007) Classification in networked data: a toolkit and a univariate case study. *J. Mach. Learn. Res.*, **8**, 935–983.

Muller, P., Devnani, S., Julius, J., Gagliardi, D. and Marzocchi, C. (2016) Annual report on European SMEs 2015/2016. London Economics, London. (Available from `https://ec.europa.eu/jrc/sites/jrcsh/files/annual_report_-_eu_smes_2015-16.pdf`.)

Organisation for Economic Co-operation and Development (2009) The impact of the global crisis on SME and entrepreneurship financing and policy responses. *Technical Report*. Centre for Entrepreneurship, SMEs and Local Development, Organisation for Economic Co-operation and Development, Paris.

Óskarsdóttir, M., Bravo, C., Sarraute, C., Vanthienen, J. and Baesens, B. (2018) The value of big data for credit scoring: enhancing financial inclusion using mobile phone data and social network analytics. *Appl. Soft Comput.*, **74**, 26–39.
Petrone, D. and Latora, V. (2018) A dynamic approach merging network theory and credit risk techniques to assess systemic risk in financial networks. *Scient. Rep.*, **8**, 55–61.
Probst, P. and Boulesteix, A. (2018) To tune or not to tune the number of trees in random forest. *J. Mach. Learn. Res.*, **18**, 1–18.
Probst, P., Wright, M. and Boulesteix, A. (2019) Hyperparameters and tuning strategies for random forest. *Data Minng Knowl, Discov.*, **9**, article e1301.
Soramäki, K., Bech, M., Arnold, J., Glass, R. and Beyeler, W. (2007) The topology of interbank payment flows. *Physica* A, **379**, 317–333.
Stahel, W. and Maechler, M. (2017) robustX: 'eXtra' / 'eXperimental' functionality for robust statistics. *R Package Version 1.2-2*. Eidgenössiche Technische Hochschule Zürich, Zürich.
Sun, X. and Xu, W. (2014) Fast implementation of DeLongs algorithm for comparing the areas under correlated receiver operating characteristic curves. *IEEE Signl Process. Lett.*, **21**, 1389–1393.
Van Vlasselaer, V., Bravo, C., Caelen, O., Eliassi-Rad, T., Akoglu, L., Snoeck, M. and Baesens, B. (2015) APATE: a novel approach for automated credit card transaction fraud detection using network-based extensions. *Decsn Supprt Syst.*, **75**, 38–48.
Weisbuch, G. and Battiston, S. (2007) From production networks to geographical economics. *J. Econ. Behav. Organizn*, **64**, 448–469.
White, S. (2015) Business population estimates for the UK and regions 2015. Department for Business, Innovation and Skills, Sheffield. (Available from https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/467443/bpe_2015_statistical_release.pdf.)