CREM
UMR CNRS 6211

Centre de Recherche en Économie et Management
Center for Research in Economics and Management

University of Caen
University of Rennes 1

Working Paper

The Relevance of Soft Information for Predicting Small Business Credit Default: Evidence from a Social Bank

Simon Cornée

CREM CNRS UMR 6211, University of Rennes 1, and CERMi, France

June 2012 - WP 2012-26

UNICAEN université de Caen Basse-Normandie

cnrs dépasser les frontières

UNIVERSITÉ DE RENNES 1 ueb

The Relevance of Soft Information for Predicting Small Business Credit Default: Evidence from a Social Bank[1]

Simon CORNEE

(Université de Rennes 1, CREM UMR CNRS 6211, and CERMi)[2]

CREM Working Paper 2012-26,

This version September 2013

ABSTRACT:

This paper analyzes whether soft information improves predictive accuracy of a bank's internal rating by better forecasting future credit defaults. We use a unique hand-collected database of 389 small loans granted by a French social bank (credit cooperative) dealing with (very) small businesses. Almost all the loans in our sample have an amount of under €250,000. These small loans are typical candidates for small business credit scoring, which is only based on hard information. Our study emphasizes the relevance of including soft information – in addition to hard information – to improve credit default prediction. Our paper contributes to the topical debate on the design of credit rating systems tailored to suit the lending technologies of different banks. This issue is of particular importance for non-conventional relationship banks such as social banks and credit cooperatives.

Keywords: Social Banking, Credit Cooperative, Credit Rating, Debt Default, Relationship Lending.

JEL codes: G21, M21.

---

[1] For their comments and suggestions, we gratefully thank Yiorgos Alexopoulos, Frédéric Basso, Carlos Borzaga, Olivier Brandouy, Charles Calomiris, Joeffrey Drouard, Pascal François, Sébastien Galanti, Georges Gallais-Hammono, Christophe Godlewski, Silvio Goglio, Frans van Helden, Panu Kalmi, Marc Labie, Eric Lamarque, Olivier L'Haridon, Frédéric Lobez, Roy Mersland, Franck Moraux, Patrick Navatte, Adrian Pop, Christian Rauch, Catherine Refait-Alexandre, Ariane Szafarz, Jean-Laurent Viviani, Lawrence Wall, Laurent Weill, Jacob Yaron, the participants at the International Symposium on Money, Banking, and Finance (June 2012), and the IV International Workshop on Cooperative Finance at Trento (June 2013).

[2] Address: Faculté des Sciences Economiques, 7 Place Hoche CS86514, 35065 RENNES Cedex, FRANCE. Email: simon.cornee@univ-rennes1.fr. Phone: +33 (0)2 23 23 38 87.

1. INTRODUCTION

The severity of asymmetric information between insiders and outsiders in the credit market renders the provision of external funding to small and atypical firms particularly challenging (Stiglitz and Weiss, 1981). These types of businesses are typically much more informationally opaque than are medium-sized enterprises and large corporations that are able to demonstrate their quality to financial markets by merely communicating audited financial statements. To address this opacity problem, financial intermediaries use a number of different lending technologies.

Financing small businesses has traditionally been a local and close-knit affair. Banks have essentially relied on tight location-based credit relationships to underwrite their small business loans (Meyer, 1998). The prominent feature of this relationship lending technology is that the credit-granting decision is based on the production of *soft information* (Berger and Udell, 2002). This type of information is gathered over time by loan officers through intense contact with firm managers (Scott, 2006), and by being embedded in the local community (Uzzi and Lancaster, 2003). Since soft information is contingent on its producers (loan officers), it is difficult to verify by third parties and subjective by nature (e.g. Stein, 2002).

The last decade has witnessed the fast development of transaction lending technologies, the leader of these technologies being Small Business Credit Scoring (henceforth, SBCS). Banks have increasingly resorted to SBCS to evaluate applicants for small loans typically under €250,000 (Berger and Frame, 2007). SBCS combines *hard information* using statistical methods to predict future credit performance. Hard information characteristically includes quantitative financial data about the firm and consumer data about the owner (Mester, 1997; Akhavein et al., 2005). The propagation of the SBCS approach is concomitant with the consolidation of the banking industry and the increasing geographic distance between small business borrowers and their bank lenders (Petersen and Rajan, 2002; Brevoort and Hannan, 2004). There exists a large body of empirical research studying the effects of the changes in lending technologies on small credit availability (Berger and Udell, 2006; Berger and Frame, 2007).

All in all, these substantial shifts in the banking industry have resulted in a greater reliance on hard information models in credit-granting decisions at the expense of qualitative approaches. In contrast, academic literature has started to pay attention to soft information – although moderately. On the one hand, the analysis has been primarily conducted at the industry level in order to explain why some banks have switched to the production of hard information while others still favor the accumulation of soft information (Berger and Udell, 2002). For example, Stein (2002) and Berger et al. (2005) show that the main type of information used by a bank to select its borrowers can be explained by its organizational architecture. Typically, small and decentralized banks with few managerial layers are at a comparative advantage in evaluating investment projects for which the information is soft by nature.

On the other hand, more recent work examines the place of soft information in micro aspects of lending, i.e. credit conditions and approval decisions. Cerqueiro et al. (2011) document that loan rate setting is not only explained by rule (statistical methods based on hard information) but that it also reflects a bank's reliance on loan officers' discretion based on soft information. Discretion in the loan-pricing process is also most important if loans and firms are small. In the same vein, discretion given to loan officers in bank lending decisions is also widespread and economically significant, especially in small banks. Interestingly, approved loans based on loan officers' judgment – derived from previously collected soft information – perform as well as other loans, *ceteris paribus* (Puri et al., 2011). Gropp et al. (2012) corroborate the previous findings and document that loan officers are even too cautious in their use of soft information in their approval decisions.

The use of soft information proves to be critical in loan approval decision-making and loan-pricing processes. Nonetheless, direct evidence on whether soft information improves the accuracy of a bank's screening is scant, especially for small loans. The objective of this paper is to fill this gap by empirically investigating the extent to which the inclusion of soft information in models of credit default prediction improves their forecast quality. Our study uses a unique hand-collected dataset including detailed information from 389 business loans granted by a French credit cooperative. Importantly, 99.48 % of the loans reported in our database are small loans exhibiting an amount of under €250,000. We show that a forecast model combining both soft and hard information outperforms the benchmark model that relies solely on hard information.

The contribution of this paper to the literature is twofold. The first contribution deals with the relevance of fitting credit rating models with soft information factors in the context of small loans.[3,4] Since the seminal works of Beaver (1966) and Altman (1968), extensive research has been carried out to prove the suitability of rating techniques based on quantitative financial data to predict corporate insolvency (Altman and Saunders, 1997). Even though the combined use of soft and hard factors in credit rating systems may intuitively lead to more accurate default prediction models,[5] very little work has been carried out so far to

[3] Like Krahnen and Weber (2001), in their normative set of "generally accepted rating principles", we posit that credit ratings are deemed to predict default. We thus consider jointly the research on the prediction of corporate bankruptcy as well as the normative and empirical literature on credit rating.
[4] Interestingly, empirical studies on internal credit rating systems document that these systems are based on statistical method or/and expert judgment (Machauer and Weber, 1998; Brunner et al., 2000; Treacy and Carey, 2000).
[5] For instance, Altman and Sabado (2007, p. 335) acknowledge that "[their] *analysis could still be improved using qualitative variables as predictors in the failure prediction model to better discriminate between small and medium enterprises*".

give empirical support to this conjecture. When this evidence proves to exist, it does not concern small business finance (Ciampi and Gordini, 2013). Using data from large German commercial banks dealing with medium-sized companies, Lehmann (2003) and Grunert et al. (2005) consider qualitative factors such as management quality or industry prospects. Both studies conclude that the combined use of financial and non-financial factors leads to better default predictions rather than the individual use of either of these factors. Our paper adds to this scarce literature by specifically focusing on small loans of under €250,000, which are the typical candidates for SBCS.

The second contribution of the paper lies in the fact we consider diversity in bank organizational structure and in lending technology by examining an unexplored but burgeoning segment of the credit market: social banks. This attribute represents an additional line of interest.[6] Social banks are committed to paying attention to the non-economic (i.e., social and/or environmental) consequences of their financial activities (Benedikter, 2011; Weber and Remer, 2011). Accordingly, they select their borrowers on a financial *and* social basis. Social banks fund micro-borrowers, small-sized enterprises, social enterprises, and cooperatives that are often ill-financed by conventional banking sector because of their strong informational opacity or their specific governance and missions (Borzaga and Defourny, 2001). Like credit cooperatives and community banks, social banks have traditionally relied on relationship lending technologies by making extensive use of soft information to assess their credit applicants (Angelini et al., 1999; De Young et al., 2004; Scott, 2004).[7] Although social banks still remain marginal in the financial landscape, they have dramatically spread

---

[6] With regards to this point, Grunert et al. (2005, p. 528) emphasize the importance of "*collecting data from different financial intermediaries*, [in order for their] *results to be tested with regards to bank size and organizational structure following Berger et al. (2005) and Stein (2002).*" A number of authors also make the case for diversity in the banking sector (e.g. Ayadi et al. 2010).

[7] Given the similarities in terms of missions, lending technology and organizational architecture between social banks, on the one hand, and credit cooperatives, on the other hand, we contend that the implications of our study can, at least to some extent, be generalized to the latter (Groenveld, 2010; Groenveld, 2011; Kalmi, 2012).

over the last decades, especially since the outbreak of the financial crisis.[8] Although they are facing crucial challenges and novel strategic choices inherent to their huge expansion, especially in terms of organizational architecture and lending practices, there is still a lack of empirical evidence that sheds light on their activity.[9] This study partially fills the gap by gaining a better understanding of social banks' credit rating systems. Moreover, this paper offers managerial implications as well as more general regulatory policy implications. Rating systems are not only used internally for credit risk management but also externally, in particular for calibrating banks' capital requirement.

The rest of the paper is organized as follows. Section 2 describes the data. Section 3 examines the extent to which soft information improves credit default prediction. Section 4 offers robustness checks. Section 5 concludes.

2. THE DATASET

2.1. INSTITUTIONAL SETTING AND SAMPLE DESCRIPTION

The hand-collected data used in this study stem from the portfolio of a French social bank. The data were extracted from the same original sample used by Cornée and Szafarz (2012). The bank under scrutiny is a credit cooperative established twenty years ago. It operates throughout the country under the supervision of the *Banque de France*, the French Central Bank. In 2010, it was made up of 27,135 members, and its total assets amounted to €288

[8] In Europe, the strong place of social banking, today's combined assets total about €21 billion according to the European Federation of Ethical and Alternative Banks (www.ethicalbankingeurope.com). Between 2007 and 2010, the asset growth rate of European social banks reached 53.41%, compared with 8.37% for mainstream banking (Global Alliance for Banking on Value (2012) cited in Cornée and Szafarz (2012)).

[9] Notable exceptions are Becchetti and Garcia (2011), Becchetti et al. (2011), Cornée et al. (2012), Cornée and Szafarz (2012).

million. The bank puts into practice traditional financial intermediation rules insofar as it bans all forms of speculative financial transactions (San-Jose et al., 2011).

The bank's simple organizational architecture involves three branches; each of these covering several French administrative regions. During the sample period, 15.5 full-time-equivalent loan officers took part in lending operations. On average, a loan officer grants 25 credits per annum. These first-line personnel play a pivotal role in the organization. They collect both soft and hard information on credit applicants by filling out a standard file. All the data, in particular the qualitative facts, are amassed *in situ* as loan officers pay a visit to most credit applicants.

The fact that the bank under study selects its debtors upon both financial and social criteria may be problematic for our study, which is typically the case if there is interference between financial and social dimensions. Cornée and Szafarz (2012) discard this possibility by showing that the bank is neither more stringent nor more flexible with social projects. However, the bank's social mission may impact its portfolio composition. For instance, its commitment to the right to credit is an incentive to serve market segments ill-financed by conventional banks such as start-ups, micro-entrepreneurs, and nonprofit organizations.

Our database consists of 389 loans and covers the 2001-2008 period. We first consider a credit-granting period stretching from 2001 to 2004 during which the loans were extended. We then scrutinize a four-year observation window from 2004 to 2008 to record potential default events. This means that we have a standardized four-year observation period for each of the 389 loans.[10] Over the 2001-2004 period, the bank grants 476 loans. Our sample therefore represents 81% of the whole population. There should not be any selection bias

[10] More precisely, the data were collected in November 2008. As a consequence, the credit-granting period and the four-year observation period stretch from January 2001 to November 2004 and from November 2004 to November 2008, respectively.

since the 19% of missing loans were not intentionally ruled out but were excluded because of information unavailability. Moreover, the representativeness of the sample betters overtime as the bank's information system improves.[11]

All borrowing firms can be regarded as *veritable* small-sized enterprises. First, their legal form is typical of small businesses: 44% are proprietorships or other forms of individual unlimited companies, 37% are private limited liability companies, and 19% are cooperatives or other legal structures of nonprofit organizations. However, the most explicit evidence of their small size is undoubtedly provided by their average turnover which amounts to about €540,000 (median: €119,000),[12] and their average number of employees which reaches 7.59 (median: 5.28). It is also noteworthy that 49% of the borrowing firms are start-ups. As argued above, this distinctive feature mostly stems from the social mission of the bank. The substantial fraction of start-ups automatically pushes the age of the borrowing firms downward (5.28 years on average).

As a corollary to the small size of the businesses under scrutiny, loan amounts are quite low. Loan size ranges from €5,000 to €300,000€, and averages at €46,900 (median: €30,520). All loans but two are under € 250,000, thereby making all of them typical candidates for SBCS. Moreover, a substantial fraction of them (43%) can be regarded as microloans according to European standards, which define all loans under €25,000 as microcredits.

As argued above, our figures are much lower than those of the existing studies that focus on medium-sized companies. Data extracted by Lehmann (2003) from a large German commercial bank report that two thirds of the companies sampled present a turnover up to €5

[11] The representativeness is 57.47% in 2001 (50 out of 87), 79.25% in 2002 (84 out of 106), 90.21% in 2003 (129 out of 143), and 90.00% in 2004 (126 out of 140).

[12] Due to data availability issue, average turnover and employees are computed on 55 firms. Average age is computed on 350 firms. All the other statistics are calculated on the whole sample (389 firms).

million. The sample used by Grunert et al. (2005) is restricted to six major German banks trading with medium-sized enterprises whose turnover amounts to a sum between €25 million and €250 million and whose minimum loan size is around €1.5 million. These figures are 10 to 25 times higher than ours.

2.2. DEFINITION OF THE VARIABLES

An observation consists of a borrowing firm's hard information rating and soft information ratings as well as its default status in the four years following loan extension. Loans are granted for maturities ranging from one to twenty years (average maturity: 7.22 years). This four-year convention set by the bank is thus somewhat arbitrary. Still, when analyzing the time period between credit granting and default occurrences, one can observe that the vast majority of defaults occur within four years following credit granting (87%).

Table 1: Definition of the variables

VARIABLES	DEFINITION
Default variable	
DEF	= 1 if a default occurs within four years following loan extension; 0 otherwise.
Hard information rating	
FIN	In-house financial rating: from 3 (best) to 1 (worst)
Soft information ratings	
MGT	Assessment of "management quality": from 3 (best) to 1 (worst).
PROJECT	Assessment of "project quality": from 3 (best) to 1 (worst).
Type of borrowing firm	
STARTUP	= 1 if the loan is extended to a start-up firm; 0 otherwise.
Year-fixed effects	
Y2001	= 1 if the loan is extended in 2001; 0 otherwise.
Y2002	= 1 if the loan is extended in 2002; 0 otherwise.
Y2003	= 1 if the loan is extended in 2003; 0 otherwise.
Y2004	= 1 if the loan is extended in 2004; 0 otherwise.

Table 1 describes the variables used in the study, broken down into four categories. First, the default variable, DEF, is a dummy variable that equals 1 if one or more of the subsequent events occur within four years following loan extension and 0 otherwise: moratorium, allowance of loan provisions, withdrawal of credit, disposition of collateral, and liquidation. This definition of default occurrence is consistent with the definition suggested by the Basel Committee on Banking Supervision (Grunert et al., 2005).

The construction of the ratings relies on the information collected by loan officers. Importantly, these ratings are given at the moment of credit-granting decisions and are not revised thereafter. The hard information rating consists of the FIN variable. This financial rating is established by loan officers according to a bank's in-house procedure. It is assigned via a backward- and forward-looking perspective by taking into account past financial statements as well as business risks and prospects. The soft information ratings correspond to the subjective judgment of loan officers for credit applicants. This judgment is split into two dimensions: MGT and PROJECT. The MGT variable provides an overall view of the management character (honesty, prudence, ethics) and capacity (experience, training and motivation). The PROJECT variable assesses qualitatively the global relevance of the investment project (Strengths, Weaknesses, Opportunities, Threats matrix).

Finally, two sets of control variables are included. First, the type of borrowing firm differentiates the loans extended to start-ups (STARTUP) from those of existing businesses. Second, Y2001, Y2002, Y2003, and Y2004 control for year effects. These dummy variables aim to capture potential changes in the global economic conjuncture that may in turn have influenced both the bank's lending strategy and its borrowers' creditworthiness. In addition to that, incomplete – and therefore excluded – files are proportionately more frequent during the year of observation (i.e., 2001). Lastly, Table 2 indicates that the distribution of default occurrences is not perfectly smooth.

Table 2: Default occurrences per annum

Year	Full sample	Full sample (%)	DEF=0	DEF=1	DEF=1(%)
2001	50	17.86%	43	7	14.00%
2002	84	36.00%	60	24	28.57%
2003	129	27.40%	98	31	24.03%
2004	126	18.52%	97	29	23.02%
Total	389	100.00%	298	91	23.89%

Table 2 displays a breakdown of default occurrences throughout the years and indicates a relatively even distribution of default, with the exception of 2001 which yielded a smaller ratio for the number of defaults over the total number of observations. In total, 23.89% of borrowers experienced a default event within four year following loan extension. This figure indicates that the sample is well-balanced in terms of default and non-default populations, and should not bias the econometric analysis (King and Zen, 2001).

Admittedly, the proportion of loans in default appears quite high according to conventional retail banking standards. To put this figure into perspective, it should be known that the percentage of defaulting loans is mechanically elevated by the fact that the bank finances about as many start-ups as existing businesses. Defaults appear to be more than twice as frequent for start-ups (32%) as for existing firms (14%). On a related note, we cannot exclude that this bank adopts a riskier lending attitude, driven by its social mission through the financing of unconventional activities. Second, the “default” denomination we use is extensive and encompasses various types of repayment issues that do not have the same degree of stringency. Cornée and Szafarz (2012) give a rough estimate of the loans in default that are eventually liquidated on the basis of out-of-sample figures from 2007. They estimate that, on average, only 3.5% of the bank’s loan portfolio results in liquidation. This represents about 2 % and 5% of the loans extended to existing companies and to start-ups, respectively. Furthermore, this social bank may adopt more prudent behavior in its risk management –

while financing riskier market segments –, given its ethical orientation. For instance, it may consider certain loans “in default” which would not necessarily be regarded as such in a conventional bank.

Table 3: Summary statistics for each rating category

	Full sample	DEF=0	DEF=1	t-test
FIN	1.97 (0.03)	2.01 (0.03)	1.84 (0.04)	0.00***
MGT	2.76 (0.02)	2.81 (0.04)	2.58 (0.05)	0.00***
PROJECT	2.27 (0.02)	2.31(0.05)	2.14 (0.04)	0.00***
Number of obs.	389	298	91	

Means are displayed and standard deviations are reported in parentheses. ***: equality between DEF=0 and DEF=1 rejected at the 1% level. **: equality rejected at the 5% level. *: equality rejected at the 10% level.

Table 3 displays the rating categories averaged within the full sample, as well as within the DEF=0 and DEF=1 subsamples. Expectedly, the means of the three rating categories are lower for defaulters than they are for non-defaulters. This is a first indication that a strong link can be established between credit ratings and default status. The first column of Table 4 confirms this robust relationship by revealing that the DEF variable exhibits significantly negative correlations with all rating categories.

Table 4: Correlation matrix

	DEF	FIN	MGT	PROJECT
DEF	1.00			
FIN	-0.15***	1.00		
MGT	-0.22***	0.19***	1.00	
PROJECT	-0.15***	0.18***	0.15***	1.00
STARTUP	0.22***	-0.05	-0.09*	-0.16***

All correlations are Spearman rank correlations. *: zero correlation rejected with $p < 10\%$, **: zero correlation rejected with $p < 5\%$, ***: zero correlation rejected with $p < 1\%$

Something else worth mentioning from Table 4 is the moderate – although significantly positive – correlation between FIN, MGT and PROJECT variables. This shows that, while converging towards the same point, the three rating variables only partially overlap, thereby revealing complementary contributions in their ability to predict default occurrences.

Finally, the significantly positive correlation between STARTUP and DEF variables confirms conventional wisdom in that start-up businesses are more likely to experience repayment issues. Given that start-ups prove to be riskier, the negative correlations between STARTUP and all rating variables seem quite normal. However, the fact that start-ups benefit from lower ratings may also reflect other problems such as informational asymmetry, which is markedly more acute for start-ups. This evidence shows the necessity of explicitly accounting for the STARTUP variable in the econometric analysis.

3. SOFT INFORMATION AND CREDIT DEFAULT PREDICTION

Here we examine the extent to which soft information improves the forecast quality of the bank's internal rating. To check whether soft information leads to a more accurate prediction of future credit events, we specify two models, each of which aims to estimate the probability of default. These models differ in the following respects: the benchmark model, or HI model, which is standard in retail banking, is fitted with only hard financial information; whereas the HISI model is enriched with soft information in addition to hard information.

In Panel A of Table 5, we run probit regressions with DEF as a dependent variable. In specification (1), the HI model includes only the financial rating as an independent variable. In the HISI model specified in column (2), the soft information ratings, MGT and PROJECT, are added. Both models control for year-specific effects. Specifications (3) and (4) replicate columns (1) and (2), respectively, except that they include the STARTUP variable.

Table 5: Forecast quality of HI and HISI models

Panel A: regression results				
	(1)	(2)	(3)	(4)
MODELS	HI	HISI	HI	HISI
FIN	-0.46***	-0.30**	-0.47***	-0.33**
	(0.154)	(0.162)	(0.161)	(0.167)
MGT		-0.56***		-0.54***
		(0.158)		(0.160)
PROJECT		-0.32**		-0.24
		(0.160)		(0.166)
STARTUP			0.61***	0.56***
			(0.148)	(0.152)
Y2002	0.53**	0.52*	0.48*	0.47*
	(0.268)	(0.276)	(0.271)	(0.279)
Y2003	0.39	0.35	0.31	0.28
	(0.255)	(0.263)	(0.259)	(0.266)
Y2004	0.35	0.46*	0.31	0.42
	(0.257)	(0.266)	(0.260)	(0.268)
CONSTANT	-0.21	1.70***	-0.46	1.25**
	(0.367)	(0.591)	(0.382)	(0.612)
Observations	389	389	389	389
McFadden's R^2	0.0308	0.0736	0.073	0.1065
LR Test	0.0111	0.0000	0.0000	0.0000

Panel B: evaluation criteria				
Area under ROC	0.6216	0.6908**	0.6804	0.1593***
Brier Score	0.1743	0.1653**	0.1666	0.7330*
Naive Brier Score	0.1790	0.1790	0.1790	0.1790

*, **, and *** indicate that the tests between 'Areas under ROC' or between Brier scores (indicating that SI model performs better than HI model) are significant at the 10%, 5%, and 1% level, respectively.

Overall, there is clear evidence to suggest a significantly negative impact for all ratings with regards to default probability. This supports the idea that both hard *and* soft information are relevant when it comes to predicting default occurrences. Interestingly, the financial rating maintains its explanatory power throughout Table 5, and does not get crowded out by non-financial ratings or by control variables. This may also be interpreted in the sense that the bank's rating system is sophisticated enough to enable us to draw general conclusions from the results. Lastly, the highly significant and negative coefficient on STARTUP reveals that the three ratings taken together are not able to fully capture all the peculiarities of start-up businesses. This confirms that the explicit inclusion of this control variable in the analysis is relevant.

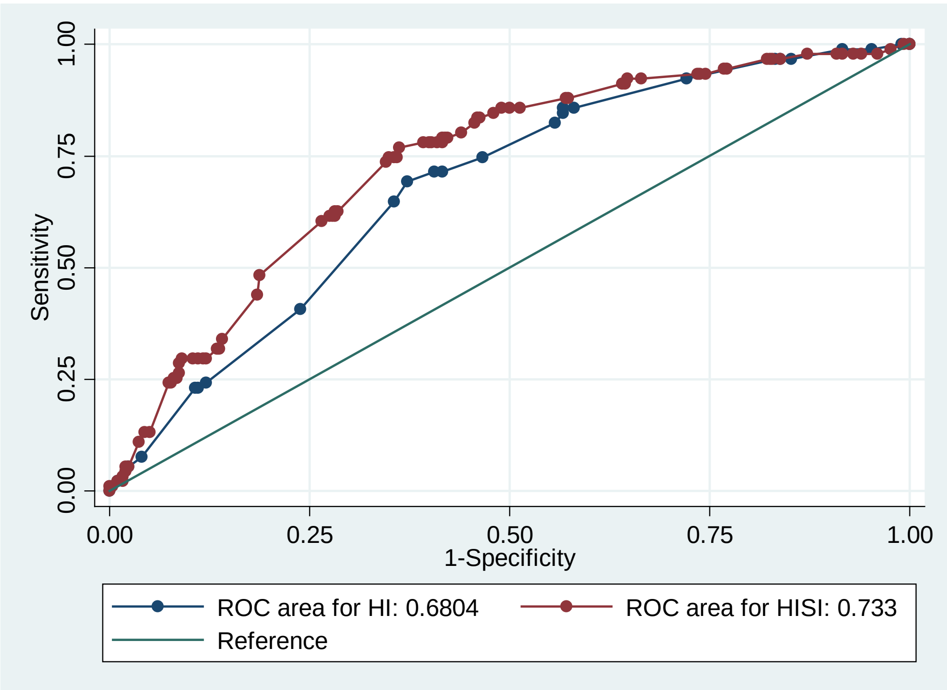
We now turn to compare the relative performance of the HI model versus the HISI model. As there is no single measure to comprehensively evaluate a classification model, two criteria were selected: the area under ROC (Receiver Operating Characteristic) curve and the Brier score. On the one hand, the ROC curve methodology is a commonly used validation technique for rating systems (Güttler, 2005; Behr and Güttler, 2007; Güttler and Krämer, 2008). In contrast with the percentage of correctly classified observations that is based on a single arbitrary set,[13] the ROC inspection that accounts for multiple thresholds avoids setting a single arbitrary threshold by obtaining the optimal combination of Type I and Type II errors over the entire range of thresholds.[14] On the other hand, we consider an alternative indicator well-known in meteorology and medical science to measure prediction accuracy: the Brier score.[15] This indicator ensures that the probit models yield predicted default occurrences better than those which could be inferred from naive forecasting. It also takes the estimated probabilities directly into consideration. This is not the case with the area under ROC, which transforms probabilities to either 0 or 1 at each threshold.

The results, reported in Panel B of Table 5 and in Figure 1, indicate that the HISI model outperforms the benchmark HI model with respect to these two evaluation criteria. Figure 1 displays the ROC curves representing each model. The area under the ROC curve can be interpreted as follows. For a given model, the further away the ROC curve from the diagonal line, i.e. the larger the area under the ROC curve, the better its performance in

[13] If a conservative threshold is set at 0.5, all the firms with a probability of default superior to 0.5 are considered as defaulters whereas the firms below this threshold are assumed to be non-defaulters.
[14] See its formal calculation and its application to credit default prediction in Lehman (2003), Güttler (2005), Behr and Güttler (2007), and Güttler and Krämer (2008).
[15] See its formal calculation in Brier (1950) and its application to credit default prediction in Grunert et al. (2005), Güttler (2005), Krämer and Güttler (2008).

predicting defaults. The diagonal line represents a random model without any discriminative power. In this case, the corresponding area is 0.5. The area for any rating model oscillates in practice between 0.5 and 1; the area of a perfect model being equal to unity. We find ROC areas of 68.04% and 73.30%, respectively, for HI model (blue curve) and HISI model (red curve) as specified in columns (3) and (4) of Table 5. This graphical impression is confirmed by a formal test of equality of the ROC areas.[16] This test rejects the null hypothesis in which ROC areas are equal for both models at the 5% level or at the 1% level, including or not including the STARTUP variable.

Figure 1: Areas under the ROC curve for SI and HI models



Examination of the Brier scores yielded by HI and HISI models corroborates the ROC inspection. A two-tailed Williams-Kloot test rejects the null hypothesis in which Brier scores are equal for both models at the 5% threshold and 10% threshold, without and with the

[16] This test is described in detail by Cleves (2002).

control variable STARTUP, respectively.[17] It is also noteworthy that both models are useful since they are always more accurate than the naive forecasting (regardless of the inclusion of the STARTUP variable).

On the whole, the HISI model clearly outperforms the standard HI model with respect to the two evaluation criteria. A mixed model that includes both hard and soft information performs better than a quantitative model that relies solely on hard information.

4. ROBUSTNESS CHECKS

4.1. BOOTSTRAP PROCEDURE

We further test the comparative performance of the two models for robustness by applying a bootstrap procedure. It is possible that the previous probit estimates are biased because of violated distributional assumptions or correlated regressors and error terms. Resorting to the bootstrap methodology enables us to address this concern (Grunert et al., 2005). We generate an empirical estimate of the sample distribution for each evaluation criterion by following three steps (Güttler, 2008). First, 389 observations are randomly drawn with replacement from the original sample. Second, the HI and HISI models that include the two sets of control variables are estimated in the same way as in columns (3) and (4) in Table 5.[18] Third, the performance of each model is evaluated using the same previous criteria (area under the ROC curve and Brier score). These three steps are independently replicated 1000 times.

On the basis of the bootstrap distribution, Panel A of Table 6 reports the bias (average difference between the bootstrap estimates and the observed statistics), its standard deviation

[17] This test is described in detail by Vinterbo and Ohno-Machado (1999).
[18] For robustness purposes, we decided to keep specifications (3) and (4) for the bootstrap procedure because they both include the two sets of control variables. We also carried out the bootstrap procedure on models (1) and (2), which give most similar results.

and the corresponding – normal or bias-corrected – confidence intervals at 95%. When the ratio of the bias over its standard deviation is below 0.25, the bias is acceptable because the random error overrides it (Efron, 1982). In our case, the aforementioned ratio almost systematically exceeds 0.25 – although to a moderate extent, since it yields values ranging from 0.23 to 0.49. For this reason, we also consider bias-corrected confidence intervals. The reported confidence intervals, whether normal or bias-corrected – are consistent with the ranking of the two models previously found.

Table 6: Bootstrapping of evaluation criteria

Panel A:

Evaluation criterion		**HI model**	**HISI model**
Area under ROC	Bootstrap mean	0.6804	0.7329
	Bias	0.0148	0.0095
	Bootstrap std. dev.	0.0298	0.0292
	95% confidence interval	(0.6228; 0.7378)	(0.6775;0.7903)
	Bias-corrected 95% confidence interval	(0.6063; 0.7259)	(0.6673;0.7793)
Brier score	Bootstrap mean	0.1666	0.1593
	Bias	-0.0025	-0.0033
	Bootstrap std. dev.	0.0107	0.0108
	95% confidence interval	(0.1456; 0.1873)	(0.1380;0.1806)
	Bias-corrected 95% confidence interval	(0.1488; 0.1874)	(0.1417;0.1812)

Panel B:

Evaluation criterion	**Prob (HISI model worse than HI model)**
Area under the ROC curve	0.10%
Brier score	0.00%

Reported values correspond to the probabilities of observing a positive/negative pairwise difference between HI model and HISI model for each evaluation criterion.

The bootstrap procedure also allows us to determine the number of cases (out of the 1000 replications) in which the HISI model exhibits better predictive ability than the HI model does. Figure 2 proposes a diagram in which the ROC area of the HISI model is on the vertical axis and that of the HI model is on the horizontal axis. With the help of this diagram, 1000 dot comparison pairs are obtained. The 45-degree line indicates model pairs of equal area under the ROC curve. Likewise, Figure 3 gives a similar graphical representation of the comparison pairs for the Brier score by plotting the bootstrap values of the HISI model against those of the benchmark HI model.

Figure 2: "Area under the ROC curve" pairs: HISI model versus HI model

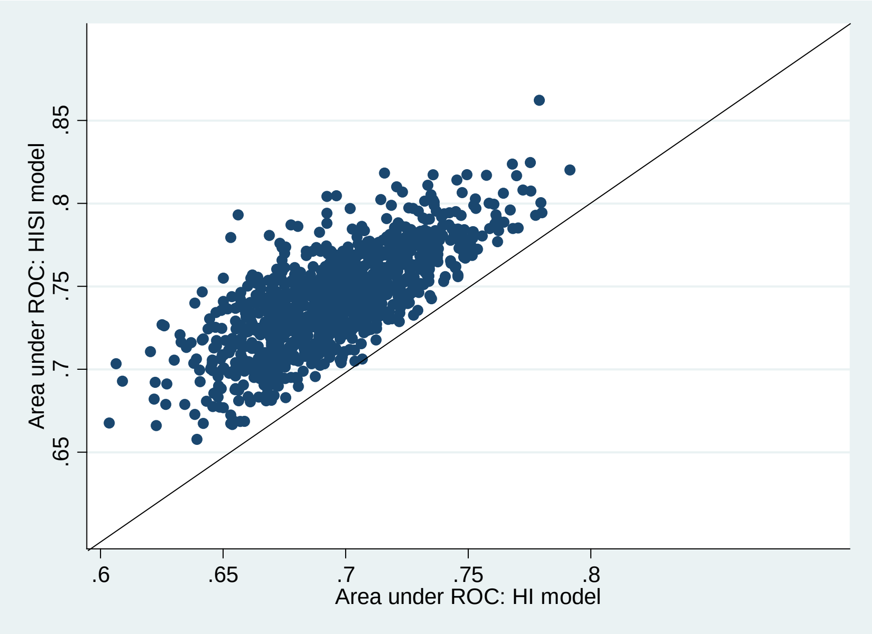
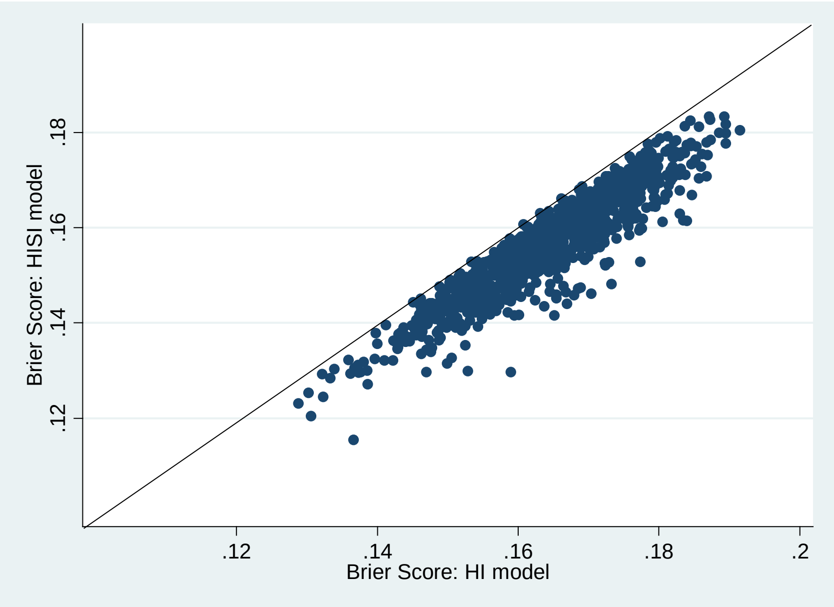


Figure 3: "Brier score" pairs: HISI model versus SI model



The probabilities corresponding to the graphical representations are reported in Panel B of Table 6. In the case of the area under ROC, the HISI model dominates the HI model in 999 out of 1000 cases. That is to say, there exists a 0.10% probability that the HISI model performs worse than the HI model. As with the Brier score, the result is even more striking since the HISI model outperforms the benchmark model in all cases.

4.2. ADDITIONAL ROBUSTNESS CHECKS

To further support the previous findings, we propose a series of four robustness checks along different directions. These checks are conducted on the HI and HISI models as specified in columns (3) and (4) of Table 5. The results are summarized in Table 7.

Table 7: Robustness checks

Robustness checks		**Model**	**# Observations**	**Area under ROC**	**Brier score**
Replacing FIN by the Banque de France's financial rating		HI	389	0.6574	0.1682
		HISI	389	0.7240***	0.1598**
Standard errors clustered at the loan officer unit		HI	389	0.6804	0.1666
		HISI	389	0.7330***	0.1593*
Removal of one year successively	**2001**	HISI	339	0.6682	0.1676
		HISI	339	0.7212***	0.1610*
	2002	HI	305	0.6777	0.1673
		HISI	305	0.7266**	15.99**
	2003	HI	260	0.6806***	0.1678
		HISI	260	0.7339	0.1623
	2004	HI	263	0.6758	0.1672
		HISI	263	0.7291***	0.1595*
Sample restricted to existing firms		HI	200	0.5850	0.1837
		HISI	200	0.6648***	0.1754**

*, **, and *** show that the tests of equality between 'Areas under ROC' or between Brier scores (indicating that the SI model performs better than the HI model) are significant at the 10%, 5%, and 1% level, respectively.

First, in both models, we replace the FIN variable by another hard information rating produced by the bank on the basis of the *Banque de France* (French Central Bank)'s guidelines. Whereas the FIN variable is used by the bank in its credit risk management, this alternative financial rating aims instead to comply with regulatory purposes. This rating, computed from financial statements of borrowers, is given on a one-to-five scale, five being the best. Our results are robust with respect to this change in the firm-specific hard information measure.

Second, we re-estimate both specifications by clustering the standard errors at the “loan officer” unit. In doing so, we account for potential differences in loan officers, who play a crucial role in the assignment of all credit ratings. Differences that do not depend on credit applicants’ intrinsic qualities such as loan officers’ experience, self-confidence, or branch remoteness from the bank’s headquarters may typically introduce unjustified variations in rating assignment. Our results do not lose any significance.

Third, we address the marginal ‘per annum’ impact because some coefficients on year dummy variables are significant in Table 5. We successively leave out one year from the sample and re-estimate both models with the three remaining years. Our results hold true.

Fourth, we rerun both models by restricting our sample to existing firms. The sample contains a large proportion of start-ups, which may somehow bias the analysis: for instance, the financial rating, which is conventionally based on past events, may not correctly account for the start-ups’ financial situations. Under this scenario, an inaccurate financial rating for start-ups would be mechanically detrimental to the HI model in its comparison to the HISI model. By only keeping existing firms, we avoid this potential bias. Interestingly, the difference in forecast quality between the two models fully remains with this subsample.

On the whole, the robustness checks confirm our previous findings on the superiority of the HI model over the HISI model. That is to say, the combined use of qualitative and quantitative information results in significantly higher predictive accuracy than the sole use of quantitative information.

5. CONCLUSION

The ability of financial intermediaries to correctly predict the default of their borrowers and to rate them directly impacts their lending behavior in terms of price and quantity. This issue is particularly critical for market segments that are the most sensitive to credit rationing by conventional banking: SMEs, micro-borrowers, nonprofit organizations, etc. Beyond their informational opacity and lack of pledgeable assets, their sensitivity to credit rationing lies in the fact that they are productive units with the lowest level of substitutability between bank loans and alternative forms of financing (Carter and Van Auken, 2006; Becchetti et al., 2011).

The empirical conclusion of the paper suggests that, in the context of a small social bank, credit-granting decisions are more accurate when they are based on both hard *and* soft information rather than when they rely solely on hard information. We therefore provide new evidence on the relevance of including soft information factors in assessing very small borrowers' creditworthiness, especially when the bank has few managerial layers and deals with an informationally opaque clientele.

Some policy implications may be drawn from the present paper. It contributes to the topical debate on the design of credit rating systems, which should at same time be robust and representative of the diversity of lending technologies. In showing the relevance of soft information to predict future credit performance, this work contributes to pave the way for credit rating systems that better espouse the peculiarities of relationship-based lending technologies. Given the importance of relationship lending in facilitating access to credit for veritable small enterprises, the regulatory framework should also be tailored accordingly (Milano, 2005; Travis, 2005). This policy implication undoubtedly applies to social banks but also to their close counterparts such as credit cooperatives, which also make an extensive use of soft information and present similar organizational forms (Ayadi, 2005; Groenveld, 2012).

It should be noted that further investigation is needed to generalize our results. Drawing global conclusions from a specific institution is always hazardous. Moreover, several issues relating to our question remain unsolved. For instance, relying on soft information mechanically empowers loan officers because they are delegated more authority in the information production process, which is vital for all financial intermediaries. Albeit monitored, loan officers still benefit from considerable managerial discretion that may in turn induce detrimental effects. For example, “social attachment” between loan officers and their borrowers may result in overlending (Uzzi and Gillespie, 1999). Conversely, the influence of loan officers on final decision making may lead to credit rationing due to discrimination. Credit officers may indeed exhibit unjustified preferences and/or stereotypes, which in turn exclude certain population segments from borrowing (*e.g.*, gender discrimination, Agier and Szafarz, 2013). Subsequently, agency problems may arise not only between lenders and borrowers, but also between loan officers and their own banks.

Another complex issue, for the future, relates to how profitability will be impacted with the inclusion of soft information in the default prediction model. The incurring costs of using soft information appear markedly multifaceted and difficult to assess. The costs derived from the agency problems mentioned above offer a good illustration (Udell, 1989). Another example is the inherent costs associated with the soft information collection, which may vary overtime. Typically, it may be high in the initial phase of the credit relationship and much lower as the bank-borrower interactions amplify (Petersen and Rajan, 1995).

In summary, this paper documents that the inclusion of soft information factors improves the forecast quality in default prediction models, when a bank handles very small or atypical businesses. Our work also enriches the scarce literature on the nascent phenomenon of social banking.

REFERENCES

Agier, I., and A. Szafarz (2013), "Subjectivity in Credit Allocation to Micro-Entrepreneurs: Evidence from Brazil," *Small Business Economics* 41 (1), 263-275.

Akhavein, J., W.S. Frame, and L.J. White (2005), "The Diffusion of Financial Innovation: An Examination of the Adoption of Small Business Credit Scoring by Large Banking Organizations," *Journal of Business* 78 (2), 577-596.

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

Altman, E.I., and G. Sabado (2007), "Modeling Credit Risk for SMEs: Evidence from the US Market," *Abacus* 43 (3), 332-357.

Altman, E.I., and A. Saunders (1997), "Credit Risk Measurement: Developments over the Last 20 Years," *Journal of Banking and Finance* 21 (11-12), 1721-1742.

Angelini, P., R. Di Salvo, and G. Ferri (1998), "Availability and Cost of Credit for Small Businesses: Customer Relationships and Credit Cooperatives," *Journal of Banking and Finance* 22 (6-8), 925-954.

Ayadi, R. (2005), "The New Basel Capital Accord and SME Financing. SMEs and the New Rating Culture," Working Paper, Centre for European Policy Studies.

Ayadi, R., D.T. Llewellyn, R.H. Schmidt, E. Arbak, and W.P. De Groen (2010), "Investigating Diversity in the Banking Sector in Europe: Key Developments, Performance and Role of Cooperative Banks," Working Paper, Centre for European Policy Studies, Brussels.

Beaver, W.H. (1966), "Financial Ratios as Predictors of Failure," *Journal of Accounting Research* 4, 7-11.

Becchetti, L., and M. Garcia (2011), "Do Collateral Theories Work in Social Banking?" *Applied Financial Economics* 21 (13), 931-947.

Becchetti, L., M. Garcia, and G. Trovato (2011), "Credit Rationing and Credit View: Empirical Evidence from Loan Data," *Journal of Money, Credit and Banking* 43 (6), 1217-1245.

Behr, P., and A. Güttler (2007), "Credit Risk Assessment and Relationship Lending: An Empirical Analysis of German Small and Medium-Size Enterprises," *Journal of Small Business Management* 45 (2), 194-213.

Benedikter, R. (2011), *Social Banking and Social Finance,* New-York: Springer.

Berger, A.N., and W.S. Frame (2007), "Small Business Credit Scoring and Credit Availability," *Journal of Small Business Management* 45 (1), 5-22.

Berger, A.N, N.H. Miller, M.A. Petersen, R.G. Rajan, and J.C. Stein (2005), "Does Function Allow Organizational Form? Evidence from the Lending Practices of Large and Small Banks," *Journal of Financial Economics* 76 (2), 237-269.

Berger, A.N., and G.F. Udell (2002), "Small Business Credit Availability and Relationship Lending: the Importance of Bank Organizational Structure," *Economic Journal* 112 (477), 34-53.

Berger, A.N. and G.F. Udell (2006), "A More Complete Conceptual Framework for SME Finance," *Journal of Banking and Finance* 30 (11), 2945-2966.

Borzaga, C., and J. Defourny (2001), *The Emergence of Social Enterprise*, London – New York: Routledge.

Brevoort, K.P., and T.H. Hannan (2004), "Commercial Lending and Distance: Evidence from Community Reinvestment Act Data," *Journal of Money, Credit and Banking* 38 (8), 1991-2012.

Brier, G.W., (1950), "Verification of Forecasts Expressed in Terms of Probability," *Monthly Weather Review* 78, 1-3.

Brunner, A., J. Krahnen, and M. Weber (2000), "Information Production in Credit Relationships: On the Role of Internal Ratings in Commercial Banking," Centre For Financial Studies Working Paper No 2000/10.

Carter, R., and H. Van Auken (2006), "Small Firm Bankruptcy," *Journal of Small Business Management* 44 (4), 493–512.

Cerqueiro, G., H. Degryse, and S. Ongena (2011), "Rules versus Discretion in Loan Rate Setting," *Journal of Financial Intermediation* 20 (4), 503-529.

Ciampi, F., and N. Gordini (2013), "Small Enterprise Default Prediction Modeling through Artificial Neural Networks: An Empirical Analysis of Italian Small Enterprises," *Journal of Small Business Management* 51(1), 23-45.

Cleves, M. (2002), "From the Help Desk: Comparing Areas under Receiver Operating Characteristic Curves from Two or More Probit or Logit Models," *Stata Journal* 3, 301-313.

Cornée, S., D. Masclet, and G. Thenet (2012), "Credit Relationships: Evidence from Experiments with Real Bankers," *Journal of Money, Credit and Banking* 44(5), 957-980.

Cornée S., and A. Szafarz (2012), “Vive La Différence: Social banks and Reciprocity in the Credit Market,” WP-CEB No 12-29.

DeYoung, R., W.C. Hunter, and G.F. Udell (2004), “The Past, Present and Probable Future of Community Banks,” *Journal of Financial Service Research* 25 (2-3), 85-133.

Efron, B. (1979), “Bootstrap Methods: Another Look at the Jackknife,” *Annals of Statistics* 7 (1), 1-26.

Groeneveld, H. (2010), “Cooperative Banks in the New Financial System,” Presentation prepared for SVP Cooperative Affairs and Governance: A Competitive Europe? The Case for Cooperatives.

Groeneveld, J.M. (2011), “Morality and Integrity in Cooperative Banking,” *Ethical Perspectives* 18 (4), 515-540.

Groeneveld, H. (2012), “The Cooperative Banking Model: Performance and Opportunities”, in: Balling, M., F. Lierman, F. Van den Spiegel, R. Ayadi, and D.T. Llewellyn (Eds.), *New Paradigms in Banking, Financial Markets and Regulations?* Vienna: Larcier, 101-130.

Gropp, R., C. Gruendl, and A. Guettler (2012), “Does Discretion in Lending Increase Bank Risk? Borrower Self-Selection and Loan Officer Capture Effects,” European Banking Center Discussion Paper No 2012-010.

Grunert, J., L. Norden, and M. Weber (2005), The Role of Non-Financial Factors in Internal Credit Ratings, *Journal of Banking and Finance* 29 (2), 509-531.

Güttler, A. (2005), “Using a Bootstrap Approach to Rate the Raters,” *Financial Markets and Portfolio Management* 19 (3), 277–295.

Kalmi, P. (2012), “Co-operative Banking,” in: Michell, J. and J. Toporowski (eds.), *Handbook of Critical Issues in Finance*, Elgar, 56-66.

King, G., and L. Zeng (2001), “Logistic Regression in Rare Events Data,” *Political Analysis* 9 (2), 137-163.

Krahnen, J.P., and M. Weber (2001), “Generally Accepted Rating Principles: A Primer,” *Journal of Banking and Finance* 25 (1), 3-23.

Krämer, W, and A. Güttler (2008), “On Comparing the Accuracy of Default Predictions in the Rating Industry,” *Empirical Economics* 34 (2), 343-356.

Lehmann, B. (2003), “Is It Worth the While? The Relevance of Qualitative Information in Credit Rating,” University of Konstanz Working Paper.

Machauer A., and M. Weber (1998), “Bank Behavior Based on Internal Credit Ratings of Borrowers,” *Journal of Banking and Finance* 22 (10-11), 1355-1383.

Mester, L.J. (1997), “What’s the Point of Credit Scoring?” *Federal Reserve Bank of Philadelphia Business Review* September/October, 3-16.

Meyer, L.H. (1998), “The Present and Future Roles of Banks in Small Business Finance,” *Journal of Banking and Finance* 22, 1109-1116.

Milano, R. (2005), “Ethical Finance in Europe Following the Basel II Agreements,” *Finance & Common Good* 25 (1), 15-17.

Petersen, M.A., and R.G. Rajan (1995), “The Effect of Credit Market Competition on Lending Relationships,” *Quarterly Journal of Economics* 110 (2), 407-443.

Petersen, M.A., and R.G. Rajan (2002), “The information Revolution and Small Business Lending: Does Distance Still Matter?” *Journal of Finance* 57 (6), 2533-2570.

Puri, M., J. Rocholl, and S. Steffen (2011), “Rules Versus Discretion in Bank Lending Decisions,” Unpublished working paper.

Scott, J.A. (2004), “Small Businesses and the Value of Community Banks, *Journal of Financial Service Research*,” 25 (2-3), 207-230.

Scott, J.A. (2006), “Loan Officer Turnover and Credit Availability for Small Firms,” *Journal of Small Business Management* 44 (4), 544-562.

Stein, J.C. (2002), “Information Production and Capital Allocation: Decentralized versus Hierarchical Firms,” *Journal of Finance* 57 (5), 1891-1921.

Stiglitz, J.E., and A. Weiss (1981), “Credit Rationing in Markets with Imperfect Information,” *American Economic Review* 71 (3), 393-410.

Travis, A. (2005), “Will Ethical Finance Survive Basel II?” *Finance & Common Good* 21 (1), 57-67.

Treacy, W.F., and M. Carey (2000), “Credit Rating Systems at Large US Banks,” *Journal of Banking and Finance*, 24 (1-2), 167-201.

Udell, G.F. (1989), “Loan Quality, Commercial Loan Review and Loan Officer Contracting,” *Journal of Banking and Finance* 13 (3), 367-382.

Uzzi, B., and J.J. Gillespie (1999), “Interfirm Relationships and the Firm’s Financial Capital Structure: the Case of the Middle Market”, *Research in the Sociology of Organizations* 16, 107–126.

Uzzi, B., and R. Lancaster (2003), “Relational Embeddedness and Learning: The case of Bank Managers and their Clients,” *Management Science*, 49 (4), 383-399.

Vinterbo, S., and L. Ohno-Machado (1999), “A Recalibration Method for Predictive Models with Dichotomous Outcomes,” in *Predictive Models in Medicine: Some Methods for Construction and Adaptation*, PhD Dissertation, Norwegian University of Science and Technology.

Weber, O., and S. Remer (2011), *Social Banks and the Future of Sustainable Finance*, London – New York: Routledge.