Quantifying Credit Gaps
Using Survey Data on Discouraged Borrowers

15th January 2025

Abstract

This paper proposes a methodology to estimate the aggregate financing needs of firms that are bankable yet discouraged from applying for a loan. Our data come from the 2018-2020 EBRD-EIB-World Bank Enterprise Survey and cover 35 emerging and developing economies. Drawing on the literature on corporate bankruptcy prediction, we develop a model with elastic net regularization to predict the outcome of loan applications. Our approach suggests that 32%-40% of discouraged firms would have had their loan application approved, signaling inefficient credit rationing. Using this information, we estimate an aggregate credit gap of 4.9%-5.6% of GDP, with significant variation across countries. SMEs account for more than two-thirds of the total, reflecting both their contribution to economic activity and the fact that they are more likely to be credit-constrained.

JEL Codes: D22, D45, E51, G21, G32
Keywords: credit rationing, discouraged borrowers, firm-level data, EMDEs

1 Introduction

Credit rationing arises from information asymmetries between borrowers and lenders, which can lead to moral hazard (Holmstrom and Tirole, 1997) or adverse selection (Stiglitz and Weiss, 1981). Both theoretical mechanisms observe that a higher interest rate reduces the borrower's stake in a project. This in turn constrains the ability of the lender to increase profits by raising interest rates. As a result, credit markets are characterized by rationing and more generally, an inefficient allocation of resources. To mitigate these market failures, Public Development Banks devote a substantial amount of resources. For instance, SME financing in 2021 accounted for €45bn of the total committed lending volume of €94.9bn for the European Investment Bank (EIB) Group. Quantifying the extent to which companies are able to obtain the finance they need is thus of first-order importance.

To this end, this paper proposes a methodology that quantifies excess demand in corporate credit markets from the bottom up, with a focus on discouraged borrowers operating in emerging and developing economies. Most empirical studies on discouraged borrowers focus on advanced economies, but discouragement may be more prevalent in developing countries (Chakravarty and Xiang, 2013). According to Levenson and Willard (2000) and Kon and Storey (2003), discouraged borrowers are creditworthy firms in need of external finance that nevertheless do not apply for a loan because they expect to be rejected and face high application costs.[1] Our methodology allows for some discouraged borrowers to be rationed for good reasons (Han et al., 2009). Providing credit to all discouraged firms is unlikely to result in an optimal allocation of resources. This paper, therefore, seeks to quantify the financing needs of firms that are discouraged from applying yet *bankable* from a credit scoring perspective.

[1]Studies of credit rationing among firms (Stiglitz and Weiss, 1981; Berger and Udell, 1992; Banerjee and Duflo, 2014; Berg, 2018) frequently focus on firms that apply for a loan.

Our methodology employs the 2018-2020 EIB-EBRD-WBG Enterprise Surveys (ES) as the main data source. Our analysis covers 23,815 firms in 35 high and middle-income economies across Europe, Asia, the Middle East and North Africa. The survey contains a detailed set of questions that measure a firm's ability to access finance. Among firms that need a loan, the survey distinguishes between firms that successfully applied for a loan, firms that had their loan application rejected, and firms that were discouraged from applying for a loan. Empirically, discouraged borrowers account for 22.2% of firms in our sample, compared to only 1.2% of rejected applicants.

The credit gap in this paper is given by the aggregate financing needs of bankable discouraged firms. To identify the set of bankable discouraged firms we first estimate a rejection model, trained to predict rejections in the sample of applicants. The Enterprise Survey enables us to construct a large set of candidate predictors. We use the elastic net (Zou and Hastie, 2005) as a regularization technique to optimize the predictive performance of the model. Among the various combinations of the Ridge and the Lasso penalty that we examine, the Lasso exhibits the lowest deviance as the measure of the mean cross-validated error. The results indicate a considerable level of discriminatory power, also when compared to bankruptcy prediction models based on financial statement information (see for instance Amendola et al. (2015); Tian et al. (2015); Tian and Yu (2017); Bai and Tian (2020)). This validates the use of survey data for our purposes. The selected covariates have the expected sign and capture meaningful economic relationships with the outcome variable. By applying the model out-of-sample we obtain rejection probabilities for the discouraged firms.

We allocate credit using different methods. First, each firm receives credit in proportion to the probability of approval. This yields an upper bound estimate of the credit gap as it assumes that there are no systematic unobserved differences between applicant and discouraged firms. Second, credit is allocated based on a misclassification cost function, which relaxes this assumption. This mechanism involves setting a rejection threshold, such that firms with a rejection probability exceeding the threshold are fully rationed, whereas firms with a rejection probability below this threshold obtain credit. Two patterns emerge from the data. First, even the least conservative method - the proportional allocation mechanism - implies that the rejection rate of discouraged firms is close to three times the in-sample rejection rate. This suggests that the average discouraged firm is less creditworthy than the average applicant. Second, the misclassification cost function suggests that between 32% and 40% of discouraged firms would have seen their loan application approved if they were to apply for a loan. Hence, even the more conservative approach indicates inefficient credit rationing.

The last step consists of aggregating the results to obtain credit gap estimates at the country and regional level. The upper bound estimates suggest a credit gap of 8.8-8.9% of GDP or USD 319-322bn for the countries covered in this study. At the same time, and not surprisingly, the estimates based on a the more conservative allocation mechanism generate a smaller credit gap amounting to 4.9-5.6% of GDP or USD 178-202bn, depending on the rejection model specification. As the survey provides information on employment in discouraged firms, we can decompose the credit gap into an SME and a corporate component. The SME component is of particular interest in our context, because they generate a large share of GDP in emerging and developing economies and play an important role in creating sufficient jobs for a growing global workforce. In addition, they generate positive externalities through innovation and technology adoption. At the same time, SMEs tend to be more opaque than corporates, and thus more prone to credit rationing. We find that SMEs account for 73% of the overall credit gap in the countries covered in this paper representing a credit gap ranging between 3.6% and 6.4% of GDP depending on the rejection model specification and the allocation mechanism.

This paper contributes to the bottom-up approach to credit gap estimation based on firm-level data (Chakraborty and Mallick, 2012; Domeher et al., 2017; IFC et al., 2017;

Lopez-de Silanes et al., 2018; Cole and Sokolyk, 2016; Corrigan et al., 2020).[2] The main methodological contribution of the paper is to link credit gap estimation to the literature on discouraged borrowers and to the literature on corporate bankruptcy prediction.

To conceptualize and measure excess demand for credit, our paper draws on the literature on discouraged borrowers (Levenson and Willard, 2000; Kon and Storey, 2003). It is thus related to studies by Brown et al. (2011); Mac an Bhaird et al. (2016); Cole and Sokolyk (2016); Rostamkalaei et al. (2020); Brown et al. (2022) and Ferrando and Mulier (2022) in that it identifies the financing needs of discouraged borrowers that are bankable from a credit screening perspective. Empirically, we use a definition of discouraged borrowers based on survey data that is in line with recent studies (Chakravarty and Xiang, 2013; Matias Gama et al., 2017; Rostamkalaei et al., 2020; Brown et al., 2022; Ferrando and Mulier, 2022; Bertrand and Mazza, 2022; Wernli and Dietrich, 2022; Cowling and Sclip, 2022; Bertrand et al., 2024).

Our modelling approach is informed by the extensive literature on corporate bankruptcy prediction (Beaver, 1966; Altman, 1968). Several studies implement probabilistic approaches to identifying firm failures and bankruptcies (Ohlson, 1980; Zavgren, 1985; Zmijewski, 1984), proposing methodologies to identify and select risk drivers notably employing machine learning based methods (Altman et al., 1977, 1994; Chava and Jarrow, 2004; Liang et al., 2016; Tobback et al., 2017; F. Mai et al., 2019; Tsai, 2009; Lin et al., 2014; Tian and Yu, 2017; G. Kou et al., 2021). In an extensive review of the literature, Bellovary et al. (2007) and Zhao et al. (2024) find that hundreds of different variables have been used as bankruptcy predictors. Balcaen and Ooghe (2006) stress that financial ratios of SMEs are often unstable and of limited usefulness in a bankruptcy prediction context. This increases the value of information from sources other than financial state-

[2] The literature has developed two approaches to credit gap measurement, each with its own purpose: (i) a macroeconomic approach, and (ii) methodologies based on firm-level data. The former is employed primarily in macroprudential contexts, such as setting countercyclical capital buffers in the context of Basel III (Drehmann and Tsatsaronis, 2014; Lang and Welz, 2018). The latter takes a bottom-up approach to quantifying structural excess demand for credit.

ments (Becchetti and Sierra, 2003; Daubie and Meskens, 2002; Ohlson, 1980; Sheppard, 1994; Zavgren, 1983; Keasey and Watson, 1987; Lussier and Pfeifer, 2001) - notably for SMEs operating in emerging and developing markets, which often lack reliable financial statement information (Altman and Sabato, 2007; Altman et al., 2010, 2016; Balcaen and Ooghe, 2006). Altman et al. (2020) notes that using non-financial variables is hardly possible in a cross-country context due to their incomparability and limited availability. This is why many studies focus on single countries when using non-financial predictors. Our study fills this gap by using a large set of candidate predictors that are comparable across countries, drawing on the Enterprise Survey.

The remainder of the paper is organized as follows: Section 2 introduces the data; Section 3 provides an account of the methodology; Section 4 presents the results. The last section concludes.

2 Data

Firm-level data come from the 2018-2020 wave of the Enterprise Surveys (ES), implemented by the European Investment Bank, the European Bank for Reconstruction and Development and the World Bank Group. Our analysis exploits data on 23,815 firms across 35 economies in Central, Eastern, South-Eastern Europe, Central Asia, the Middle East, and North Africa. Table 1 provides a list of the countries covered in the analysis. To facilitate comparisons across countries and regions, we group them based on geographic proximity. The Enterprise Survey covers a representative sample of an economy's formal, non-agricultural private sector. It includes a broad range of business environment topics, notably access to finance, corruption, infrastructure, crime, competition, investment decisions as well as firm performance. Enterprise Surveys involve face-to-face interviews with business owners and top managers and are designed to represent the business environment as experienced by firms. The samples are stratified by size, sector, and geography. Large firms are over-sampled to allow for inference at a reasonable sample size.[3] As the sampling probability differs across firms, we use sampling weights during the aggregation process.

The goal of our analysis is to identify the set of firms that are creditworthy, yet rationed. To this end, we can draw on a detailed set of widely used questions (Popov and Udell, 2012; Gorodnichenko and Schnitzer, 2013) that measure a firm's ability to access finance. Of particular interest are firms that need a loan, but are discouraged from applying (Freel et al., 2012; Kon and Storey, 2003). We start by identifying firms that desire bank loans. These are composed of firms that applied for a loan, i.e. that answer affirmatively to question K16: "Did the establishment apply for any loans or lines of credit in the last fiscal year?". Firms that did not apply are then asked question K17: "What was the main reason the establishment did not apply for any line of credit or loan in the last fiscal year?". Firms that answer "Interest rates are not favorable"; "Collateral requirements are too high"; "Size of loan and maturity are insufficient"; or "Did not think it would be approved" also need a loan, but are discouraged from applying. This empirical definition is in line with recent studies (Chakravarty and Xiang, 2013; Matias Gama et al., 2017; Rostamkalaei et al., 2020; Brown et al., 2022; Ferrando and Mulier, 2022; Bertrand and Mazza, 2022; Wernli and Dietrich, 2022; Cowling and Sclip, 2022; Bertrand et al., 2024). Discouraged firms are credit-constrained (Jappelli, 1990; Nucci et al., 2020; Pietrovito and Rancan, 2024), but they are not the only firms that are credit-constrained. In addition, firms that applied for a loan, but had their loan application rejected are also credit constrained.

In total, approximately 38% of firms in the economies covered by the Enterprise Surveys desired bank credit during the last financial year. As Table 2 shows, 16% of firms did actually apply for a loan[4], whereas 22% were discouraged from doing so. The vast majority of credit-constrained firms are discouraged from applying for a loan, as

[3] For more details, see `https://www.enterprisesurveys.org/en/methodology`.

[4] These firms can have their loan application accepted or rejected.

only 1.2% of companies have their loan application rejected. Empirically, rejections do not appear salient, but in our context, they are important to gauge the creditworthiness of discouraged firms.

The level of financial intermediation exhibits considerable heterogeneity across the countries and regions in our sample. A high share of applicants indicates active financial intermediation. A high share of discouraged firms, on the other hand, points to potentially substantial credit rationing. According to Table 2, the share of applicants ranges from 7% in the SN to 27% in WB. This reflects the low application rates in Egypt and the high weight of Egypt in the SN average. The share of discouraged firms ranges from 11% in WB to 36-37% in EN and TUR. The regions differ substantially also in the ratio of applicants to discouraged firms. CEE and WB have the highest ratio, whereas SN has the lowest ratio of applicants to discouraged firms across all countries. This gives a first indication of a potentially large credit gap in SN.

Following the existing literature, we assess the creditworthiness of discouraged borrowers based on a pool of thirty-four potential predictors of firm failures, drawing from a large set of candidates available in the ES. A full list of the variables and their definition is provided in Table 3.

In addition, our methodology requires information on the flow of credit to non-financial corporations. This information is not readily available and needs to be estimated. We do, however, have data on the stock of credit to non-financial corporations that come from the Financial Soundness Indicators compiled by the International Monetary Fund (IMF). For CEE countries, we use data on non-financial corporate (NFC) credit from the European Central Bank (ECB). In cases where these are not available, we resort to data from the IMF FAS database or to the central bank of the country. Figure 1 plots the level of NFC credit relative to GDP by country and region. With the exceptions of Lebanon and Jordan, the level of NFC credit is well below the euro area average of 41% (derived from 2019 ECB data). As the Enterprise Survey does not cover agriculture, we adjust the stock of NFC credit with the share of value added generated by the industrial and services sector, obtained from the World Bank.

To translate information on the stock of NFC credit into an estimate of the flow, we exploit information on the maturity structure of loans that is available in the 2018-2020 wave of the Enterprise Survey. Specifically, the question BMk10 asks respondents for the original maturity of the last outstanding loan. Figure 2 presents average maturity by country and region, which ranges from 0.8 years in Tunisia to 4.5 years in Albania. Though both countries have a comparable stock of NFC credit of around 21-22% of GDP, the shorter maturity in Tunisia implies that a greater proportion of the credit stock is rolled over, translating into a higher gross flow of credit.

We link the stock of NFC credit with the maturity distribution as follows:

$$credit\ flow_{i,t} = st_i\ credit_{i,t-1} + (1 - st_i)\ \frac{credit_{i,t-1}}{maturity_i^{lt}} + \Delta credit_{i,t,t-1}, \quad (1)$$

where the proportion of loans with an original maturity of one year or less is given by st_i, which, on average, applies to around 30% of loans.[5] The stock of NFC credit, adjusted for the share of value added in industry and services, is given by $credit_{i,t}$, whereas $maturity_i^{lt}$ denotes the average maturity of long-term loans, i.e. loans with an original maturity exceeding one year. Finally, $\Delta credit_{i,t,t-1}$ represents net credit growth in nominal terms, computed as the difference in the stock of two consecutive years.

Our analysis also makes use of selected macro-financial fundamentals. We use data on GDP per capita from the World Economic Outlook database of the IMF. The output gap is defined as the difference between GDP growth in 2018 and the average GDP growth between 2010 and 2019, also based on the IMF WEO database. The political instability/absence of violence dimension of the Worldwide Governance Indicators serves

[5] Some countries have a high share of non-responses to question BMk10. To account for this, we compute $st_i = (1 - nr_i) st_i^{raw} + nr_i\ st$, where nr_i is the share of non-responses in country i, and st the unconditional sample average. We proceed analogously with $maturity_i^{lt}$.

as a proxy for institutional quality. Data on the capital adequacy ratio of the banking system, the loan-to-deposit ratio, the ratio of non-performing loans to gross loans and the return on assets likewise come from the IMF Financial Soundness Indicators, and in case they are not available from national central banks.

We also use the Financial Soundness Indicators to calibrate the misclassification cost function. Specifically, we use information contained in the balance sheet and income statements to compute the average lending rate, the average deposit rate and the ratio of non-interest expense to total assets. The Credit Loss Database by Ong et al. (2023) provides information on loan write-off rates at the country-year level derived from information in BankScope.

3 Methodology

3.1 Allocating Credit to Discouraged Firms

This paper estimates the volume of additional credit that would be required to meet the needs of firms, taking into account their creditworthiness. Though we do not have data on default probabilities, the Enterprise Survey provides detailed information on firms' ability to access external finance, including the outcome of loan applications. This information helps us identify the firms in need of a loan. Firms not requiring a loan are not relevant to our analysis. Firms needing a loan are either applicants or discouraged firms. Applicants are subject to a screening mechanism and thus fall into two categories depending on whether they are approved or rejected.

The subsequent analysis is predicated on the following, stylized sequential screening mechanism for $P(rejected_j|applied_j = 1)$, which is the probability of firm j seeing its loan application rejected, conditional on having applied for a loan. Based on its profitability targets, risk policies, strategic planning, as well as its cost structure, notably its cost of capital, a financial institution sets its risk appetite. This risk appetite determines a threshold probability of default $\tilde{q}$, above which a bank rejects loan applications. If firm j applies for a loan, a bank assesses the firm's riskiness by estimating its probability of default q_j. Firm j's loan application is rejected or accepted depending on whether q_j exceeds the selection threshold $\tilde{q}$.

$$P(rejected_j|applied_j) = 1 - P(approved_j|applied_j) = P(q_j \leq \tilde{q}|applied_j). \tag{2}$$

3.1.1 Identifying Candidate Predictors

Considering the link to the default probability shown in Equation 2, we need to identify variables that predict whether a loan application is approved or rejected. To this end, we draw on the extensive literature on corporate bankruptcy prediction. Although early studies can be traced back to the 1930s (Bellovary et al., 2007), statistical tools to assess firm fragility and failures have emerged toward the end of the 1960s. Beaver (1966) in a univariate and Altman (1968) in a multivariate framework are seminal works in discriminant analysis. Subsequently, several studies have implemented probabilistic approaches to identifying firm failures and bankruptcies (Ohlson, 1980; Zavgren, 1985; Zmijewski, 1984) and developed methodologies to identify and select their risk drivers (Altman et al., 1977, 1994; Chava and Jarrow, 2004; Liang et al., 2016; Tobback et al., 2017; F. Mai et al., 2019; Tsai, 2009; Lin et al., 2014; Tian and Yu, 2017; G. Kou et al., 2021).

Selecting an efficient combination of predictors is an important issue in modelling the probability of firm failure and therefore in determining a reliable model (Du Jardin, 2009). However, there is no unequivocal evidence if favour of a narrow set of predictors. In an extensive review of the literature, Bellovary et al. (2007) and Zhao et al. (2024) find that hundreds of different variables have been used, ranging from accounting-based financial ratios over market-based financial indicators to variables based on survey data. Variables capturing the macroeconomic context have also been considered useful (Fernández-Gámez et al., 2020). In addition, Balcaen and Ooghe (2006) find little consensus on which financial variables are the best for discriminating between failing and viable firms. Furthermore, it is stressed that financial ratios of SMEs are often unstable and of limited usefulness in a bankruptcy prediction context, which increases the value of information from sources other than financial statements.

Not surprisingly, several authors advise to include non-accounting or qualitative indicators in failure models (Becchetti and Sierra, 2003; Daubie and Meskens, 2002; Ohlson, 1980; Sheppard, 1994; Zavgren, 1983; Keasey and Watson, 1987; Lussier and Pfeifer, 2001). Focusing on non-financial statements information is particularly appropriate when studying SMEs - notably in emerging and developing markets - which often lack reliable financial statement information (Altman and Sabato, 2007; Altman et al., 2010, 2016; Balcaen and Ooghe, 2006). Against this background, the literature has examined a large set of variables (Chen et al., 2011; Appiah and Chizema, 2016; Altman et al., 2010; Ahmad, 2019; Amendola et al., 2015; Bai and Tian, 2020; Tian and Yu, 2017; Putra et al., 2020; Taffler, 1983, 1984; Hill et al., 1996; Daubie and Meskens, 2002; Wang and Guedes, 2024; Becchetti and Sierra, 2003; Flagg et al., 1991; Sheppard, 1994). Along similar lines, our study captures the opacity of a firm with the number of employees, the firm's age, whether it has audited financial statements, or a prior relationship with a bank. We measure the firm's sophistication with indicators for whether it has a formal business strategy, is innovative, and whether it is active in international markets. Additional dimensions of interest include the availability of collateral, the firm's legal status, as well as its ownership structure. Altman et al. (2020) notes that using non-financial variables is hardly possible in a cross-country context due to their incomparability and limited availability. As a result, many studies that rely on non-financial predictors focus on single countries. Our study fills this gap by using a large set of candidate predictors that are comparable across countries, drawing on the Enterprise Survey.

3.1.2 Regularization to Prevent Overfitting

With a comprehensive set of candidate predictors at hand, a rejection model can be estimated (in-sample) and used to predict rejection probabilities for discouraged firms (out-of-sample). The prediction error measures how well a model performs out of sample and it can be decomposed into three components: (i) the squared estimation bias, and (ii) the variance of the model, (iii) the inherent noise in the data that cannot be predicted by any model.

The bias of the rejection model decreases with model complexity. A more complex model can better capture the underlying patterns of applicants, hence reducing bias. Conversely, the variance of the rejection model increases with model complexity. As the rejection model becomes more complex (e.g., by adding more predictors), it becomes more sensitive to the variations in the training data of applicants. This increased sensitivity causes the predictions to vary more when the model is trained on different subsets of applicants, leading to higher variance.

Considering this bias-variance trade-off, regularized regression as a prediction technique offers several advantages over using all predictors in a regression. By introducing a penalty term that selectively shrinks some coefficients and/or eliminates others, regularized regression effectively reduces the complexity of the model, preventing overfitting and improving its ability to generalize to new data. Furthermore, it improves the numerical stability of the regression coefficients, especially in the presence of multicollinearity, where predictors are highly correlated.[6]

We employ the elastic net (Zou and Hastie, 2005) as a regularization technique when fitting a binomial logit. The penalty term in the elastic net combines the ℓ_1 penalty of

[6]Several studies of corporate bankruptcy prediction use regularized regression for these advantages. See, for example, Tian et al. (2015); Amendola et al. (2015); G. Wang et al. (2016); Pereira et al. (2016); Tian and Yu (2017); Mselmi et al. (2017); Paraschiv et al. (2021); Zhao et al. (2024). In the context of credit scoring models, Albanesi and Vamossy (2024) shows that machine learning techniques are superior to traditional approaches, whereby the former reduce borrower misclassification rates by more than 40%. The advantage is greatest for smaller and more opaque counterparts, thus improving access to finance for potential new borrowers.

Lasso (Tibshirani, 1996), which encourages sparsity by setting some coefficients exactly to zero, and the ℓ_2 penalty of Ridge (Tikhonov, 1963; Hoerl and Kennard, 1970), which discourages large coefficients and thus stabilizes the estimation of correlated predictors. As opposed to starting with the Lasso or the Ridge directly, we prefer using the elastic net to refrain from an ex-ante imposition of sparsity or density on our data (Giannone et al., 2021). On the one hand, we want to allow for selecting a small set of predictors with the highest predictive power (*Lasso*). On the other hand, we acknowledge that all our predictors might be important, but then we want to shrink their impact to prevent overfitting (*Ridge*).

Following Friedman et al. (2010), the coefficient estimates are then given by minimizing the penalized negative binomial log-likelihood

$$\widehat{\beta}(\alpha, \lambda) = \arg\min_{\beta} \left\{ -\ell(\beta_0, \beta) \, + \, \lambda \, P_\alpha(\beta) \right\}, \tag{3}$$

where the binomial log-likelihood is

$$\ell\left(\beta_0, \beta\right) = \frac{1}{N} \sum_{i=1}^{N} y_i \cdot \left(\beta_0 + x_i^\top \beta\right) - \log\left(1 + e^{\left(\beta_0 + x_i^\top \beta\right)}\right), \tag{4}$$

and the elastic net penalty is

$$P_\alpha(\beta) = (1-\alpha)\frac{1}{2}\|\beta\|_{\ell_2}^2 + \alpha\|\beta\|_{\ell_1} = \sum_{j=1}^{p}\left[\frac{1}{2}(1-\alpha)\beta_j^2 + \alpha\left|\beta_j\right|\right], \tag{5}$$

with N and p denoting the sample size and the number of predictors, respectively. Ultimately, the estimates rely on tuning parameters that control the type and degree of penalization: α is the mixing parameter that controls the balance between the Lasso ($\alpha = 1$) and the Ridge ($\alpha = 0$) penalties, and λ controls the amount of regularization, i.e. how much to shrink the coefficients. For α, we consider the following grid $\{0, 0.25, 0.5, 0.75, 1\}$.

3.1.3 Model Selection and Predictive Performance

Next, we employ k-fold cross-validation to select the penalty term λ for a given value of α.[7] Our dataset exhibits class imbalance, with fewer *positive* events (loan rejections) compared to *negative* events (loan approvals). This imbalance can lead to sensitivity in k-fold cross-validation because different folds may have different distributions of rejected firms, resulting in significant variability in the selected predictors and overall model performance.

This issue can be particularly relevant for the common 5-fold or 10-fold choices. To mitigate this sensitivity, we set k equal to the sample size n in our cross-validation approach, thus effectively using leave-one-out cross-validation (LOOCV). This approach ensures that each data point is used as both a training and test point, providing a consistent and reliable evaluation of our rejection model; the iteration continues until each data point is used exactly once as a test point. Consequently, regardless of the random seed used in cross-validation, we obtain stable predictor selection and model performance.

As for assessing model performance, we employ a two-step approach: *within* and *across* models. To facilitate the exposition, we assume that there are two sets of candidate predictors, x_1 and x_2.

In the first step, within each elastic net logit rejection model, we compare different mixing parameters ($\alpha \in \{0 = \textit{Ridge}, 0.25, 0.5, 0.75, 1 = \textit{Lasso}\}$) using deviance as a measure of the mean cross-validated error. For example, when we compare the models with the same predictors x_1 but with different α values using deviance, we are essentially tuning the hyperparameters within a consistent modeling framework to find the best fit to our data. This step is crucial for model selection and regularization tuning.

[7] When selecting penalty parameters, Abadie and Kasy (2019) highlight the importance of using data-driven procedures like cross-validation, as opposed to information criteria like AIC.

In the second step, our goal shifts to evaluating the predictive power and thus selecting the best predicting models, one for each of x_1 and x_2, by comparing their Area Under the ROC (Receiver Operating Characteristic) Curve (AUC) values.[8] The AUC provides a robust, threshold-independent measure of a model's ability to discriminate between classes, making it ideal for assessing overall predictive performance.

This two-step approach leverages the strengths of both metrics: deviance for fine-tuning the model's parameters and AUC for evaluating predictive accuracy across different variable sets. Hence, we ensure that our final model is both well-tuned and highly predictive.

3.1.4 Allocating Credit in Proportion to the Probability of Approval

To estimate the credit gap, we need to assess the creditworthiness of discouraged firms as in Wernli and Dietrich (2022), Rostamkalaei et al. (2020) and Han et al. (2009). By applying the rejection model out of sample, we obtain rejection probabilities for the discouraged firms. However, an estimated rejection probability does not directly indicate whether a firm should obtain credit or not. In this study, we employ two different but complementary approaches to allocating credit. The first approach allocates credit in proportion to the probability of approval. The main advantage of this approach is that it refrains from setting a rejection threshold.

Specifically, the proportional allocation mechanism assumes that credit is allocated in proportion to the estimated approval probability of firm j in country i, $\mathbb{P}(\widehat{approved}_{ij})$. As a result, a firm with an rejection probability of 1% obtains 99% of the desired credit whereas a firm with an rejection probability of 10% receives only 90% of the desired credit. This works, because we are dealing with a sample of firms that represent a large number of firms in the economy. Thus, for a sample firm that represents 20 firms in

[8] The ROC curve plots the true positive rate against the false positive rate for each level of the rejection threshold.

the economy and has a rejection probability of 10%, 2 firms in the economy would be denied credit, whereas 18 would have their loan application approved.

It is important to note that proportional allocation assumes that there are no systematic differences between applicant and discouraged firms that are unobservable to the econometrician. This assumption is likely to be violated. Owing to the information asymmetries that give rise to credit rationing in the first place, we have reason to believe that discouraged firms are on average less creditworthy in ways that are not represented by the scoring model. Thus, the proportional allocation mechanism should be viewed as providing an upper bound estimate for the measurement of a credit gap.

3.1.5 Allocating Credit Based on Misclassification Costs

The second approach to allocating credit in our study uses a misclassification cost function, which enables us to relax the assumption that there are no systematic unobserved differences between applicant and discouraged firms. Specifically, credit is allocated to firms with a rejection probability below a rejection threshold $\tilde{p}$, whereas those with a rejection probability above it are fully rationed. To derive the rejection threshold, we use a misclassification cost function analogous to those in the literature on corporate bankruptcy prediction (Altman et al., 1977; Frydman et al., 1985; Koh, 1992):

$$MC(\tilde{p}) = RC \cdot FNR(\tilde{p}) \cdot P(rejected) + FPR(\tilde{p}) \cdot P(approved). \tag{6}$$

The misclassification cost function trades off the cost of lending to a firm that is not creditworthy against the opportunity cost of not lending to a good firm. The problem is framed such that banks seek to identify firms that are not creditworthy. Two types of errors can occur: First, high quality firms can be classified as not creditworthy, a false positive or type I error. Hence, $FPR(\tilde{p})$ gives the ratio of high quality firms not obtaining credit relative to the total number of high quality firms. Second, low quality firms can be classified as creditworthy, a false negative or type II error. Thus, $FNR(\tilde{p})$ denotes the false negative rate, i.e. the ratio of low quality companies obtaining credit relative to the total number of low quality firms. Both the false negative rate and the false positive rate are functions of the rejections threshold $\tilde{p}$. As the rejection threshold declines, the false positive rate increases whereas the false negative rate decreases.

Relative to frequently used measures such as overall accuracy, the misclassification cost function has two advantages. First, it takes into account the differing base rates of rejected and approved loan applications. As the share of approved loan applications exceeds that of rejected applications, the likelihood of committing a type II error is much higher than that of a type I error. Second, it allows for the economic consequences of the two errors to differ, as RC denotes the cost of extending credit to a firm that is not creditworthy relative to the opportunity cost of not lending to a good firm. Specifically, it may be more costly to lend to a firm that is not creditworthy than to forego a limited margin on lending to a good firm.

A drawback of the misclassification cost function is that it does not lead to a unique credit allocation but a range of allocations that is conditional on the ancilliary parameter RC. If RC is sufficiently small, all discouraged firms will obtain credit, which is unlikely to lead to a better allocation of credit. Conversely, if RC is sufficiently high, the credit gap vanishes, which from our viewpoint is also unlikely to lead to a better allocation. In what follows, we seek to identify a set of plausible values for RC.

As indicated, proportional allocation yields an upper bound estimate of the credit gap. Though this allocation is derived without reference to RC, it is consistent with a certain level of RC. Relative costs smaller than this level can be ruled out as implausible, and we do not need to consider the corresponding credit gap estimates. To construct a lower bound for RC we propose to calibrate RC such that it corresponds to an allocation that may be considered conservative. What we refer to as relative costs are de facto given by the expected returns of committing a type I and type II error, respectively, hence $RC = |ER_{II}/ER_{I}|$, where ER_i is given by

$$ER_j = q_j \cdot rr + (1 - q_j)(1 + i_m) - 1, \quad j = \{I, II\}, \tag{7}$$

where rr denotes the recovery rate and i_m the intermediation margin. Equation 2 shows that the probability of rejection and the probability of default are related. In Equation 7, q_I denotes the average probability of default for accepted firms, whereas q_{II} refers to the average default probability for rejected firms.

3.2 From Firm-Level Data to Country-Level Aggregates

So far, the analysis has focused on the individual firm. The next step is to aggregate the experiences of the individual firms to the country-level credit gap. To this end, we propose the following definition:

$$credit\ gap_i = \sum_{j \in discouraged} w_{ij}\ \mathbb{P}(\widehat{approved}_{ij})\ \widehat{volume}_{ij}, \tag{8}$$

where w_{ij} is the survey weight of firm j in country i and $\mathbb{P}(\widehat{approved}_{ij})$ corresponds to the model implied probability of approval. The term $\widehat{volume}_{ij}$ indicates the desired loan volume of the discouraged firms.

The Enterprise Survey does not ask discouraged firms for the loan amount that they would desire in case they could obtain a loan. As the likelihood of approval, this quantity is unknown and therefore needs to be approximated. To obtain a proxy, we assume that discouraged firms desire the same volume of credit per worker as the successful applicants. This strategy is feasible, as we have information on employment in both discouraged firms and successful applicants. Moreover, the Enterprise Survey asks respondents with an outstanding loan for the total balance at the time of the interview. Unfortunately, this variable has many missing values. We therefore use the aggregate volume of NFC credit scaled by the total employment of successful applicants.

This yields the following expression for the credit gap in country i:

$$credit\ gap_i = credit\ flow_i \frac{\sum_{j \in discouraged} w_{ij}\ \mathbb{P}(\widehat{approved_{ij}})\ emp_{ij}}{\sum_{k \in applied} w_{ik}\ \mathbb{1}(approved_{ik})\ emp_{ik}}, \tag{9}$$

where emp_{ij} is the full-time equivalent employment of discouraged firm j in country i, emp_{ik} is the full-time equivalent employment of firm k in country i with an approved loan application and $credit\ flow_i$ is defined in Equation 1. As Equation 9 shows, the credit gap is increasing in the total employment of discouraged firms that according to the scoring model would be eligible for credit in case they had applied. Conversely, the credit gap is decreasing in the total employment of successful loan applicants. Perhaps counter-intuitively, the credit gap is increasing in the total credit flow. This follows from linking the desired credit volume of discouraged firms to what could be referred to as a measure of leverage in successful applicants. At this stage, it is straightforward to decompose the credit gap into an SME and a corporate component

$$credit\ gap_i^{SME} = credit\ flow_i \frac{\sum_{j \in discouraged} w_{ij}\ \mathbb{P}(\widehat{approved_{ij}})\ \mathbb{1}(SME_{ij})\ emp_{ij}}{\sum_{k \in applied} w_{ik}\ \mathbb{1}(approved_{ik})\ emp_{ik}}, \tag{10}$$

where $\mathbb{1}(SME_{ij})$ takes value of one if and only if firm j is an SME.

When a threshold approach is applied to discriminate between creditworthy and not creditworthy discouraged firms, Equation 9 needs to be modified; as a result $\mathbb{P}(\widehat{approved_{ij}})$ is substituted with $\mathbb{1}(\widehat{approved_{ij}})$, where $\mathbb{1}(\widehat{approved_{ij}})$ equals one if and only if the probability of rejection is below the threshold probability $\tilde{p}$.

4 Results

4.1 Rejection Model

The objective of the model described in section 3.1 is to identify a set of variables that predict the outcome of a loan application and to use it out of sample to assess the creditworthiness of discouraged firms. In principle, we are able to generate a large number of candidate predictors from the Enterprise Survey. However, owing to missing values the training sample shrinks as the number of regressors increases. Therefore, we apply the elastic net to a model with 34 candidate predictors. Table 3 provides variable definitions, while Table 4 presents the corresponding summary statistics for both applicant and discouraged firms.

Table 5 presents results for 10 different specifications that differ along two dimensions. The first dimension concerns the set of candidate predictors. The survey-sector-country (SSC) model includes 34 survey variables, 6 sector fixed effects and 35 country fixed effects. Sector fixed effects control for the firm's industry: manufacturing, construction, retail trade, wholesale trade, hotel and restaurants, and a residual category. The survey-sector-macro (SSM) model features the same survey variables and sector fixed effects, but instead of country-fixed effects includes 7 macro-financial variables. The macroeconomic variables considered are log GDP per capita as a measure of economic development, the output gap as a proxy for the cyclical position of the economy, and political stability as measured by the Worldwide Governance Indicators to account for institutional quality. Banking sector characteristics include the capital adequacy ratio, the loan to deposit ratio, the non-performing loan ratio and the return on assets. SSM is a more parsimonious model as it includes only 47 candidate predictors compared to 75 for SSC. The second dimension concerns the relative weight of the Lasso penalty. Table 5 presents results for $\alpha \in \{0, 0.25, 0.5, 0.75, 1\}$, where $\alpha = 1$ corresponds to the Lasso penalty.

The model with country fixed effects outperforms the model with macrofinancial variables. We first identify for both SSC and SSM the specification that exhibits the lowest deviance. As Table 5 shows, all specifications exhibit a lower deviance than the unpenalized version, demonstrating the benefits of regularization. For both SSC and SSM, the deviance declines with α, such that Lasso emerges as the preferred specification. Our data therefore validate the sparsity assumption underlying Lasso, as the SSC Lasso has 36 parameters compared to 75 for the SSC Ridge. To discriminate between SSC and SSM, we turn to the AUC. The AUC of SSC equals 0.81, compared to 0.78 for SSM. This result indicates a considerable level of discriminatory power, also when compared to bankruptcy prediction models based on financial statement information (see for instance Tian and Yu (2017), Amendola et al. (2015), Bai and Tian (2020), Tian et al. (2015)) and validates the use of survey data for our purposes.

The enhanced predictive power of Lasso does not came at the cost of weaker interpretability. Table 5 ranks the predictors by the absolute value of the coefficient as a measure of importance and presents their sign. The SSC Lasso includes 12 survey variables that all have the expected sign. The three most important predictors are having an overdraft facility, having purchased fixed assets, and expecting a decline in sales over the next year. Firms with an overdraft facility are less likely to have their loan application rejected. These firms already have a relationship with a bank that mitigates information asymmetries. In addition, they may have been granted an overdraft facility because they passed a creditworthiness assessment. Next, firms that have purchased fixed assets are also less likely to have their loan application rejected. This may partly reflect reverse causality from bank loans to purchases of fixed assets. Here, it is important to note that the goal of the exercise is prediction, not to uncover the true parameter values. In contrast, firms that expect a decline in sales over the next year are more likely to have their loan application rejected. This is also intuitive as a decline in sales may compromise the firm's ability to service its debt.

Our sample includes countries that are very different in terms of economic development and structure. The models in Table 5 allow for country-specific heterogeneity via country-fixed effect and macrofinancial variables. The link between the survey variables and the outcome of the loan application, on the other hand, is assumed to apply uniformly across countries. Table 6 relaxes this assumption by allowing the coefficients on the survey variables to vary across the five regions. Within regions, countries are institutionally and culturally comparatively more homogeneous than across. For example, all economies in CEE are EU member states, and all countries in the SN share a common language.

Allowing for the coefficients on the survey variables to vary across regions yields the highest AUC of the models that we consider. To economize on space we restrict attention to the Lasso penalty, which in any case has yielded the best performing models so far. Table 6 shows that the survey-sector-country fixed effects model with regional interactions (RSSC) is more complex. It has 113 parameters in total compared to 36 for SSC. The most important predictors of SSC are selected in most regions, though their relative importance may vary. The overdraft facility, for instance, is ranked 1st in CEE but only 10th in CA. At the same time, some predictors that are not selected in SSC matter in one or more regions. Female ownership, for example, predicts a higher probability of rejection, but only in EN and SN. When it comes to performance, RSSC yields an AUC of 0.86, which is a considerable improvement over the 0.81 of SSC, as Figure 3 illustrates. But because SSC performs adequately yet is much simpler, we proceed by reporting results for both SSC and RSSC.

To illustrate the discriminatory power of the model, Figure 4A and Figure 4B present the distributions of the probability of a rejected loan application for firms whose loan application has been approved and for firms whose loan application has been rejected. Allowing for regional interactions, firms with an approved loan application have an average probability of rejection 6.7%, compared to 19.6% for firms with a rejected loan application. Though the latter figure may appear low at first glance, it follows from the low frequency of rejections in the training data.

Discouraged firms have on average a higher model-implied probability of rejection than firms with an approved loan application. Figure 4C presents the results of the out-of-sample prediction for discouraged firms. Allowing for regional interactions, the average probability of rejection for discouraged firms equals 14.4%, which is more than twice as high as the 6.7% of approved applicants. This suggests that, based on observables, discouraged firms are on average less creditworthy than successful applicants. Table 4 provides insights as to why this is the case. On average, discouraged firms have higher readings of variables that predict a rejected loan application. For example, discouraged firms are less likely to have an established bank relationship as reflected in an overdraft facility. Whereas 53% of applicants have an overdraft facility, this applies to only 32% of discouraged firms. In addition, they are less likely to own collateral in the form of real estate, as only 68% of discouraged firms own their building compared to 73% of applicants. Furthermore, discouraged firms are less internationalized than applicants as reflected in a lower share of import license applications.

4.2 Proportional Allocation

We first present our upper bound estimate of the credit gap, with each firm obtaining credit in proportion to its probability of approval. Hence, we aggregate the firm-level outcomes using Equation 9. The proportional allocation approach results in a credit gap of 8.9% of GDP or USD 322bn using the SSC specification, and of 8.8% of GDP or USD 319bn with the RSSC specification.

Table 8 presents the estimates by country and region. SN has the highest credit gap, both in absolute terms (USD 108bn) and relative to GDP (19.8%). The regional aggregate is driven by large credit gaps in Egypt (about USD 42bn) and Morocco (USD

27-28bn). Relative to GDP, Jordan (42%-43%) and Lebanon (23%-24%) also have large credit gaps. This appears counter-intuitive, given the large stock of credit to non-financial corporations in both countries (see Figure 1). However, in the case of Lebanon, the survey was implemented during a period in the second half of 2019, when the crisis affecting Lebanon intensified, resulting in a high share of discouraged companies. Turkey also has a credit gap of roughly USD 97bn according to both models, but that accounts for only 12.5% of GDP. Turkey is similar to Lebanon in that it has a fairly developed financial system, as reflected in a comparatively high share of credit to GDP. At the same time, macroeconomic conditions were deteriorating while the survey was in the field. For various reasons, the other regions have comparatively small credit gaps. In EN and, to a certain extent, CA the on average lower credit gaps are the result of high rejection probabilities that limit the amount of credit allocated to discouraged firms. Moreover, the credit flows are somewhat lower than in other regions. Ukraine stands out with the largest credit gap setting at 16%. This indicates the still relatively limited financial intermediation in the country, with a large share of state-owned banks dominating the market. In addition, a high degree of dollarization and the first Russian invasion of eastern Ukraine had a deleterious effect on asset quality. Moldova also shows an elevated credit gap, which can be explained by the shallow credit market in the country.

SMEs account for more than 70% of the overall credit gap in the countries covered in this paper. Columns (5) to (10) of Table 8 provide detailed results on the SME credit gap. SSC and RSSC again yield very similar results, with a credit gap of about USD 227bn, which accounts for 6.3% of GDP. At 13.6% (SSC) and 13.5% (RSSC), SN has the highest SME credit gap relative to GDP, whereas Turkey has the highest gap in nominal terms (roughly USD 78bn). Columns (9) and (10) of Table 8 yields the percentage of the total credit gap that is due to SMEs. In all regions with the exception of EN, SMEs account for more than 60% of the credit gap. This reflects both their contribution to economic activity and the fact that they are more likely to be credit-constrained. The lower gap on SMEs in the EN region may underscore the still significant presence of large corporate organisations that is a legacy of the Soviet era. It is not surprising that the regional aggregate is largely driven by Belarus and Ukraine, whilst Moldova, Georgia and Azerbaijan are more in line with the other regions.

At 6.3% of GDP, even our upper bound estimate of the SME credit gap is much smaller than the 19% estimated by IFC et al. (2017). This reflects differences in methodology. IFC et al. (2017) use the credit intensity of MSMEs in ten advanced benchmark economies to derive potential demand by MSMEs in emerging and developing countries.[9] But these levels of credit can only be sustained in an advanced economy context, characterized by the corresponding institutions and high levels of physical and human capital. Our study, by contrast, draws on the credit intensity of successful applicants to derive the potential demand of bankable discouraged firms located in the same country. By construction, these firms face the same operating environment as the benchmark firms. It is therefore not surprising that adding the credit gap of 8.9% of GDP to the stock of outstanding credit of 22% of GDP amounts to less than the euro area average of 41% of GDP.

4.3 Threshold Identification

The next step is to allocate credit based on rejection thresholds. To this end, we evaluate the misclassification cost function (see Equation 6) for a relative cost $RC \in \{0, ...15\}$. Table 7 presents results of this exercise for both the SSC and the RSSC model. The Column labelled *Rejection Threshold* presents the threshold probability associated with the given level of RC for each model. As expected, the threshold probability declines as the relative cost of lending to a rejected firm increases. At a relative cost of 0, (almost) all

[9] In addition, they impute - via a regression approach - the outstanding stock of MSMEs credit for those countries where data are not available.

firms obtain credit, resulting in a false negative rate exceeding 99% and a false positive rate of 0%. As the rejection threshold declines, so does the false negative rate, whereas the false positive rate increases. Table 7 shows predicted rejections both in-sample and out-of-sample. The Column *In-Sample* yields the share of firms in the training sample of the model that would have had their loan application rejected, given the threshold probability. As lending becomes more risky, the in-sample rejection rate increases. The Column *Out-of-Sample* yields the share of discouraged firms that would have had their loan application rejected conditional on the threshold probability. The on average lower creditworthiness of discouraged firms shows up in predicted rejection rates that at any given level of relative cost are consistently higher than those predicted for the training sample of actual applicants.

This approach yields a large set of allocations, not all of which are equally plausible. As indicated above, allocating credit in proportion to the probability of approval can be viewed as providing an upper bound estimate of the credit gap. Though this result has been derived without reference to the misclassification cost function, we can map it into Table 7 by backing out the implied RC level that generates a credit gap of the same size. It turns out that proportional allocation is consistent with $RC = 2.5$ for both the SSC and RSSC specifications. Hence, there is no reason to consider allocations with a lower relative misclassification cost. Mapping the proportional allocation into a misclassification cost also allows interpreting the credit gap in terms of implied rejections for discouraged firms. The SSC model allocates credit to about 82% of discouraged firms, whereas RSSC allocates credit to about 78% of firms - see Figure 4C. This compares to an approval rate of about 92% for applicants. These findings suggests that discouraged firms, if they were to apply for a loan, would face a rejection rate close to three times higher than applicants, even if we do not allow for systematic unobserved differences between applicants and discouraged firms.

The lower bound, on the other hand, occurs at the relative cost at which the credit gap vanishes. As we cannot derive a lower bound, we instead propose to calibrate the misclassification cost function in Equation 7 to arrive at an allocation that may be considered conservative. For all countries with available data we obtain the average write-off rate in the years 2017, 2018, and 2019, which amounts to 1.4%. In conjunction with a recovery rate of 40% , this yields a probability of default of 2.33%. We proxy the intermediation margin i_m as the difference between the average lending rate, the average deposit rate and the non-interest expense margin. For 2017, 2018, and 2019, we obtain an estimate of i_m that is equal to 2.9%. This translates into an opportunity cost of not lending to a good firm of 1.4%. Rejected firms, do not obtain credit and therefore we need to make an assumption on the average probability of default for these firms. Specifically, we assume an average probability of default of 30%. Hence, they are very risky, but not necessarily en route to bankruptcy. We believe that this assumption is conservative, as 22.2% of firms in the sample are discouraged (see Table 2). If more than 30% of these firms would go bankrupt every year this would have serious implications for firm dynamics in these economies. Given a recovery rate of 40% and i_m as above, lending to a rejected firm would generate a loss of 16%, which in turn yields $RC = 16/1.4 \approx 11$. Table 7 shows that the corresponding allocation is indeed rather conservative. SSC would allocate credit to only 32.6% of discouraged firms, whereas RSSC would allocate credit to 39.8% of discouraged firms. This point is also depicted on the ROC curve in Figure 3. Table 7 also shows that the resulting allocation is relatively robust to the assumption on the probability of default for rejected firms. Assuming a default probability of 35% would result in a relative cost of approximately 13, which has, however, no implications for the allocation and hence the credit gap.[10]

[10]Youden's J-statistic was proposed as a way to summarise the performance of diagnostic tests mainly in the medical field (Youden, 1950) and it was often used to identify a cut-off point on the ROC curve (Schisterman et al., 2005). Recently, it has also been used for threshold selection in the context of bankruptcy prediction models (Stankova, 2023). The J-statistic is defined as $J = sensitivity + specificty - 1$ and maximizes the distance between the ROC curve and the 45 degree line. It can be mapped into

4.4 A More Conservative Estimate

A more conservative scenario can be derived by applying the threshold parameterised in section 4.3. In this case, the results are comparatively more sensitive to the rejection model. As Table 9 shows, SSC yields a credit gap of USD 178bn, which corresponds to 4.9% of GDP or 55% of the proportional allocation estimate. By comparison, RSSC yields a credit gap of USD 202bn, which accounts for 5.6% of GDP or 63% of the proportional allocation estimate.

At the regional level, comparing the threshold allocation results against the proportional allocation estimates yields interesting findings. The lower rejection threshold affects SN firms in particular. The SSC estimate shrinks by 13.3 percentage points of GDP, while the RSSC estimate declines by 10 percentage points. In contrast, the results in Turkey are mildly affected by the tighter specification, indicating higher firm quality. The SSC estimate declines by 0.9 percentage points of GDP, whereas the RSSC estimate shrinks by only 0.1 percentage points. At the same time, a significant decline is also visible in EN, CA and CEE. In GDP terms, however, these changes are less pronounced than in the SN, because the proportional allocation estimates were smaller to begin with. Depending on the model specification and the region, the credit gap shrinks by between 1.2% and 4.3% of GDP. In relative terms, however, the threshold based allocations generate on average 50% smaller credit gap estimates than proportional allocation. At the same time, the relative position of each country within a region is largely maintained when comparing the threshold-based results to proportional allocation.

This scenario however is not the end of the line. It should be acknowledged that many more - even more conservative - estimates are compatible with the threshold selection mechanism devised in section 3.1.5. Figure 5 illustrates the sensitivity of

the misclassification cost function as $J = 1 - (FNR + FPR)$, which shows that the J-statistic gives equal weight to the false positive and the false negative rate. Hence, the J-statistic corresponds to a misclassification cost function with $RC = P(approved)/P(rejected)$, and with $P(approved) = 92.3\%$. It follows that $RC \approx 11$. Incidentally, this yields the same allocation as the calibration of the misclassification cost.

credit gap estimates with regard to RC. Figure 5A shows results corresponding to Table 7; highlighting the more stringent estimate implied by $RC = 11$. Zooming out and looking at a wider range of RCs in Figure 5B, it is clear that there is ample room to obtain even more conservative estimates. However, even by assuming that rejected firms default with a probability of 100%, it is not possible to generate a RC exceeding 42. Higher relative misclassification costs require a lower numerator than we have derived empirically. The numerator represents a country average, and as such there is dispersion around the average. Hence, there exist countries with a lower level of i_m. However, it is less plausible to think of a lower i_m for all countries. As a result, we view the allocations in Figure 5B more as they are technical solutions that can be obtained out of the optimisation problem. However they may lie outside the limits of conservative, yet plausible, boundaries.

In addition, Figure 5B shows that even at elevated RC levels, the aggregate credit does not vanish in full, suggesting that rationing is a robust feature of credit markets. The SSC credit gap declines to below 1% of GDP for $RC >= 95$ and further decreases to 0.17% of GDP for $RC >= 104$. At the same time, the RSSC specification continues to yield a persistent credit gap even at $RC = 125$. In fact, the RSCC credit gap decreases to less than 1% of GDP only for $RC >= 216$. All in all, the results clearly show that the credit gap does not disappear even if we assume that the rejected firms are on average significantly more risky. Figure 4C provides intuition as to why this is the case. In the end, a significant share of discouraged firms has a probability of rejection that is comparable to that of successful applicants.

5 Conclusion

This paper proposes a methodology to quantify credit gaps based on firm-level data. Having an idea of the size of the potential credit gaps can inform the design of policy measures that seek to reduce them. We define the credit gap as the financing needs of firms that are discouraged from applying for a loan yet bankable according to our methodology.

To identify the set of bankable discouraged firms and allocate credit we estimate loan application models, trained to predict rejections in the sample of loan applicants and informed by the extensive literature on corporate bankruptcy prediction. These models are then deployed out-of-sample to obtain rejection probabilities for the discouraged firms. Credit is allocated in different ways either proportionally to the model implied probability of approval or by inferring a cut-off threshold in the distribution of rejection probabilities of discouraged firms, thus allowing for systematic unobserved differences between applicants and discouraged firms.

We find that discouraged firms are on average less creditworthy than applicants. Yet, many of them are observationally similar to successful applicants. The share of discouraged firms that would obtain credit depends on the rejection model and the credit allocation mechanism. It ranges from a minimum of 32.6% to an upper bound of 78%. This points to inefficient credit rationing in either case.

Our results suggest a credit gap ranging from 4.9% to 8.9% of GDP in aggregate for the countries covered in this study. SMEs account for about 70% of the overall credit gap. This reflects both their contribution to economic activity and the fact that they are more likely to be credit-constrained.

Eliminating the credit gap would bring the average stock of NFC credit in the region for the period 2018-2020 to 27%, employing the conservative estimate, and no more than 31% of regional GDP, making use of the upper bound estimates. Thus, even with the credit gap closed the volume of credit would still remain well below the euro area average of 41% over the same period. This could reflect the on average lower levels of economic and financial development in the countries studied (Beck et al., 2006; Love,

2003) as well as limitations of the overall institutional framework (Demirgüç-Kunt and Maksimovic, 1998; Beck et al., 2005).

Closing the credit gaps requires a multi-year perspective and efforts from multiple actors. Larger gaps, above all in the SME segment, call for long-term funding support and an efficient interest rate pass-through to firms. Risk-sharing products can help decrease banks' risk aversion and ease the collateral requirements imposed on firms. Finally yet importantly, strengthening financial literacy (Cowling and Sclip, 2022), improving the information environment (Bertrand and Mazza, 2022) and fostering client trust in the banking sector (Koomson et al., 2023) can reduce the likelihood that firms in need of a loan are discouraged from applying.

References

ABADIE, A. AND M. KASY (2019): “Choosing among Regularized Estimators in Empirical Economics: The Risk of Machine Learning,” *Review of Economics and Statistics*, 101, 743–762.

AHMAD, A. (2019): “What factors discriminate reorganized and delisted distressed firms: Evidence from Malaysia,” *Finance Research Letters*, 29, 50–56.

ALBANESI, S. AND D. F. VAMOSSY (2024): “Credit Scores: Performance and Equity,” *NBER Working Paper Series*, 32917.

ALTMAN, E. (1968): “Financial Ratios, Discriminant Analysis and Prediction of Corporate Bankruptcy,” *Journal of Finance*, 23, 589–609.

ALTMAN, E., M. IWANICZ-DROZDOWSKA, E. LAITINEN, AND A. SUVAS (2016): “Financial and Nonfinancial Variables as Long-horizon Predictors of Bankruptcy,” *Journal of Credit Risk*, 12, 49–78.

——— (2020): “A Race for Long Horizon Bankruptcy Prediction,” *Applied Economics*, 52, 1–20.

ALTMAN, E. AND G. SABATO (2007): “Modelling Credit Risk for SMEs: Evidence from the U.S. Market,” *Abacus*, 43, 332–357.

ALTMAN, E., G. SABATO, AND N. WILSON (2010): “The Value of Non-financial Information in Small and Medium-sized Enterprise Risk Management,” *Journal of Credit Risk*, 6, 1–33.

ALTMAN, E. I., R. G. HALDEMAN, AND P. NARAYANAN (1977): “Zeta Analysis: a New Model to Identify Bankruptcy Risk of Corporations,” *Journal of Banking and Finance*, 1, 29–54.

ALTMAN, E. I., G. MARCO, AND F. VARETTO (1994): “Corporate Distress Diagnosis: Comparisons using Linear Discriminant Analysis and Neural Networks,” *Journal of Banking and Finance*, 18, 505–529.

AMENDOLA, A., M. RESTAINO, AND L. SENSINI (2015): “An analysis of the determinants of financial distress in Italy: A competing risks approach,” *International Review of Economics and Finance*, 37, 33–41.

APPIAH, K. AND A. CHIZEMA (2016): “The impact of board quality and nomination committee on corporate bankruptcy,” *Advances in Accounting*, 35, 75–81.

BAI, Q. AND S. TIAN (2020): “Innovate or die: Corporate innovation and bankruptcy forecasts,” *Journal of Empirical Finance*, 59, 88–108.

BALCAEN, S. AND H. OOGHE (2006): "35 Years of Studies on Business Failure: An Overview of the Classic Statistical Methodologies and Their Related Problems," *The British Accounting Review*, 38, 63–93.

BANERJEE, A. V. AND E. DUFLO (2014): "Do Firms Want to Borrow More? Testing Credit Constraints Using a Directed Lending Program," *Review of Economic Studies*, 81, 572–607.

BEAVER, W. (1966): "Financial Ratios as Predictors of Failure. Empirical Research in Accounting, Selected Studies," *Journal of Accounting Research*, 4, 71–111.

BECCHETTI, L. AND J. SIERRA (2003): "Bankruptcy risk and productive efficiency in manufacturing firms," *Journal of Banking and Finance*, 27, 2099–2120.

BECK, T., A. DEMIRGÜÇ-KUNT, L. LAEVEN, AND V. MAKSIMOVIC (2006): "The Determinants of Financing Obstacles," *Journal of International Money and Finance*, 25, 932–952.

BECK, T., A. DEMIRGÜÇ-KUNT, AND V. MAKSIMOVIC (2005): "Financial and Legal Constraints to Growth: Does Firm Size Matter?" *Journal of Finance*, 60, 137–177.

BELLOVARY, J., D. GIACOMINO, AND M. AKERS (2007): "A Review of Bankruptcy prediction Studies: 1980 to Present," *Journal of Financial Education*, 1, 3–41.

BERG, T. (2018): "Got Rejected? Real Effects of Not Getting a Loan," *Review of Financial Studies*, 31, 4912–4957.

BERGER, A. N. AND G. F. UDELL (1992): "Some Evidence on the Empirical Significance of Credit Rationing," *Journal of Political Economy*, 100, 1047–1077.

BERTRAND, J., P.-O. KLEIN, AND F. PASIOURAS (2024): "National culture of secrecy and firms' access to credit," *Journal of Banking and Finance*, 168, 107277.

BERTRAND, J. AND P. MAZZA (2022): "Borrowers' Discouragement and Creditor Information," *International Review of Law and Economics*, 72.

BROWN, M., S. ONGENA, A. POPOV, AND P. YEŞIN (2011): "Who needs credit and who gets credit in Eastern Europe?" *Economic Policy*, 26, 93–130.

BROWN, R., J. M. LIÑARES-ZEGARRA, AND J. O. WILSON (2022): "Innovation and Borrower Discouragement in SMEs," *Small Business Economics*, 59, 1489–1517.

CHAKRABORTY, A. AND R. MALLICK (2012): "Credit Gap in Small Businesses: Some New Evidence," *International Journal of Business*, 17, 66–80.

CHAKRAVARTY, S. AND M. XIANG (2013): "The international evidence on discouraged small businesses," *Journal of Empirical Finance*, 20, 63–82.

CHAVA, S. AND R. A. JARROW (2004): "Bankruptcy Prediction with Industry Effects," *Journal of Finance*, 8, 537–569.

CHEN, N., B. RIBEIRO, A. VIEIRA, AND J. DUARTE, J.AND NEVES (2011): "A genetic algorithm-based approach to cost-sensitive bankruptcy prediction," *Expert Systems with Applications*, 38, 12939–12945.

COLE, R. AND T. SOKOLYK (2016): "Who Needs Credit and Who Gets Credit? Evidence From the Surveys of Small Business Finances," *Journal of Financial Stability*, 24, 40–60.

CORRIGAN, E., C. O'TOOLE, AND R. SLAYMAKER (2020): "Credit Demand in the Irish Mortgage Market: What is the Gap and Could Public Lending Help?" *ESRI Working Paper No. 671*.

COWLING, M. AND A. SCLIP (2022): "Dynamic Discouraged Borrowers," *British Journal of Management*.

DAUBIE, M. AND N. MESKENS (2002): "Business Failure Prediction: A Review and Analysis of the Literature," *In: Zopounidis, C. (eds) New Trends in Banking Management. Contributions to Management Science. Physica, Heidelberg.*

DEMIRGÜÇ-KUNT, A. AND V. MAKSIMOVIC (1998): "Law, Finance, and Firm Growth," *Journal of Finance*, 53, 2107–2137.

DOMEHER, D., G. MUSAH, AND N. HASSAN (2017): "Inter-Sectoral Differences in the SME Financing Gap: Evidence from Selected Sectors in Ghana," *Journal of African Business*, 18, 194–220.

DREHMANN, M. AND K. TSATSARONIS (2014): "The Credit-to-GDP Gap and Countercyclical Capital Buffers: Questions and Answers," *BIS Quarterly Review*.

DU JARDIN, P. (2009): "Bankruptcy Prediction Models: How to Choose the Most Relevant Variables?" *Bankers, Markets and Investors*, 98, 39–46.

F. MAI, F., S. TIAN, C. LEE, AND L. MA (2019): "Deep learning models for bankruptcy prediction using textual disclosures," *European Journal of Operational Research*, 274, 743–758.

FERNÁNDEZ-GÁMEZ, M. Á., J. A. CAMPOS SORIA, J. A. C. SANTOS, AND D. ALAMINOS (2020): "European country heterogeneity in financial distress prediction: An empirical analysis with macroeconomic and regulatory factors," *Economic Modelling*, 88, 398–407.

FERRANDO, A. AND K. MULIER (2022): "The Real Effects of Credit Constraints: Evidence From Discouraged Borrowers," *Journal of Corporate Finance*, 73, 102171.

FLAGG, J. C., G. A. GIROUX, AND C. E. WIGGINS JR. (1991): "Predicting corporate bankruptcy using failing firms," *Review of Financial Economics*, 1.

FREEL, M., S. CARTER, S. TAGG, AND C. MASON (2012): "The Latent Demand for Bank Debt: Characterizing "Discouraged Borrowers"," *Small Business Economics*, 38, 399–418.

FRIEDMAN, J., T. HASTIE, AND R. TIBSHIRANI (2010): "Regularization Paths for Generalized Linear Models via Coordinate Descent," *Journal of Statistical Software*, 33, 1.

FRYDMAN, H., E. I. ALTMAN, AND D.-L. KAO (1985): "Introducing recursive partitioning for financial classification: the case of financial distress," *Journal of Finance*, 40, 269–291.

G. KOU, G., Y. XU, F. PENG, Y.AND SHEN, Y. CHEN, K. CHANG, AND S. KOU (2021): "Bankruptcy prediction for SMEs using transactional data and two-stage multiobjective feature selection," *Decision Support Systems*, 140.

G. WANG, G., G. CHEN, AND Y. CHU (2016): "A new random subspace method incorporating sentiment and textual information for financial distress prediction," *Electronic Commerce Research and Applications*, 29, 30–49.

GIANNONE, D., M. LENZA, AND G. E. PRIMICERI (2021): "Economic Predictions With Big Data: The Illusion of Sparsity," *Econometrica*, 89, 2409–2437.

GORODNICHENKO, Y. AND M. SCHNITZER (2013): "Financial Constraints and Innovation: Why Poor Countries Don't Catch Up," *Journal of the European Economic Association*, 11, 1115–1152.

HAN, L., S. FRASER, AND D. J. STOREY (2009): "Are Good or Bad Borrowers Discouraged From Applying for Loans? Evidence from US Small Business Credit Markets," *Journal of Banking and Finance*, 33, 415–424.

HILL, N. T., S. E. PERRY, AND S. ANDES (1996): "Evaluating Firms In Financial Distress: An Event History Analysis," *Journal of Applied Business Research*, 12, 60–71.

HOERL, A. E. AND R. W. KENNARD (1970): "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, 12, 55–67.

HOLMSTROM, B. AND J. TIROLE (1997): "Financial Intermediation, Loanable Funds, and the Real Sector," *Quarterly Journal of Economics*, 112, 663–691.

IFC, WORLD BANK, AND SME FINANCE FORUM (2017): "MSME Finance Gap: Assessment of the Shortfalls and Opportunities in Financing Small, Micro and Medium Enterprises in Emerging Markets," *IFC Study*.

JAPPELLI, T. (1990): "Who is credit constrained in the US economy?" *Quarterly Journal of Economics*, 105, 219 – 234.

KEASEY, K. AND R. WATSON (1987): "Non-Financial Symptoms and the Prediction of Small Company Failure: A Test of Argenti's Hypotheses," *Journal of Management*, 14, 335–354.

KOH, H. C. (1992): "The sensitivity of optimal cutoff points to misclassification costs of type I and type II errors in the going-concern prediction context." *Journal of Business Finance and Accounting*, 19.

KON, Y. AND D. J. STOREY (2003): "A Theory of Discouraged Borrowers," *Small Business Economics*, 21, 37–49.

KOOMSON, I., P. KOOMSON, AND A. ABDUL-MUMUNI (2023): "Trust in banks, financial inclusion and the mediating role of borrower discouragement," *International Review of Economics and Finance*, 88, 1418–1431.

LANG, J. H. AND P. WELZ (2018): "Semi-Structural Credit Gap Estimation," *ECB Working Paper*, No. 2194.

LEVENSON, A. R. AND K. L. WILLARD (2000): "Do Firms Get the Financing They Want? Measuring Credit Rationing Experienced by Small Businesses in the U.S." *Small Business Economics*, 14, 83–94.

LIANG, D., C. LU, C. TSAI, AND G. SHIH (2016): "Financial ratios and corporate governance indicators in bankruptcy prediction: A comprehensive study," *European Journal of Operational Research*, 252, 561–572.

LIN, F., D. LIANG, C. YEH, AND J. HUANG (2014): "Novel feature selection methods to financial distress prediction," *Expert Systems with Applications*, 41, 2472–2483.

LOPEZ-DE SILANES, F., J. MCCAHERY, D. SCHOENMAKER, AND D. STANISIC (2018): "Estimating the Financing Gaps of SMEs," *Journal of Corporate Finance Research*, 12, 1–54.

LOVE, I. (2003): "Financial Development and Financing Constraints: International Evidence From the Structural Investment Model," *Review of Financial Studies*, 16, 765–791.

LUSSIER, R. N. AND S. PFEIFER (2001): "A Cross-National Prediction Model for Business Success," *Journal of Small Business Management*, 39, 228 – 239.

MAC AN BHAIRD, C., J. SANCHEZ VIDAL, AND B. LUCEYC (2016): "Discouraged Borrowers: Evidence for Eurozone SMEs," *Journal of International Financial Markets, Institutions and Money*, 44, 46–55.

MATIAS GAMA, A. P., F. DIAS DUARTE, AND J. P. ESPERANÇA (2017): "Why discouraged borrowers exist? An empirical (re)examination from less developed countries," *Emerging Markets Review*, 33, 19–42.

MSELMI, N., A. LAHIANI, AND T. HAMZA (2017): "Financial distress prediction: The case of French small and medium-sized firms," *International Review of Financial Analysis*, 50, 67–80.

NUCCI, F., F. PIETROVITO, AND A. F. POZZOLO (2020): "Imports and credit rationing: a firm-level investigation," *The World Economy*, 44, 3141–3167.

OHLSON, J. (1980): "Financial Ratios and the Probabilistic Prediction of Bankruptcy," *Journal of Accounting Research*, 18, 109–131.

ONG, L. L., C. SCHMIEDER, AND M. WEI (2023): "Insights into credit loss rates: a global database," *Bank for International Settlements working paper*, 1101.

PARASCHIV, F., M. SCHMID, AND R. R. WAHLSTRØM (2021): "Bankruptcy Prediction of Privately Held Smes Using Feature Selection Methods," *Available at SSRN: https://ssrn.com/abstract=4588767 or http://dx.doi.org/10.2139/ssrn.4588767*.

PEREIRA, J., M. BASTO, AND A. DA SILVA (2016): "The logistic lasso and ridge regression in predicting corporate failure," *Procedia Economics and Finance*, 39, 634–641.

PIETROVITO, F. AND M. RANCAN (2024): "Credit rationing and sustainable activities: a firm-level investigation," *International Review of Economics and Finance*.

POPOV, A. AND G. F. UDELL (2012): "Cross-Border Banking, Credit Access, and the Financial Crisis," *Journal of International Economics*, 87, 147–161.

PUTRA, S., B. JOSHI, J. REDI, AND A. BOZZON (2020): "A credit scoring model for smes based on social media data," *In: Bielikova, M., Mikkonen, T., Pautasso, C. (eds) Web Engineering. ICWE 2020. Lecture Notes in Computer Science(), vol 12128. Springer, Cham*, 113–129.

ROSTAMKALAEI, A., M. NITANI, AND A. RIDING (2020): "Borrower Discouragement: The Role of Informal Turndowns," *Small Business Economics*, 54, 173–188.

SCHISTERMAN, E., N. PERKINS, A. LIU, AND H. BONDELL (2005): "Optimal cut-point and its corresponding Youden Index to discriminate individuals using pooled blood samples," *Epidemiology*, 16, 73–81.

SHEPPARD, P. (1994): "Strategy and bankruptcy: An exploration into organizational death," *Journal of Management*, 20, 795–833.

STANKOVA, M. (2023): "Threshold Moving Approach with Logit Models for Bankruptcy Prediction," *Computational Economics*, 61, 1251–1272.

STIGLITZ, J. E. AND A. WEISS (1981): "Credit Rationing in Markets with Imperfect Information," *American Economic Review*, 71, 393–410.

TAFFLER, R. (1983): "The Assessment of Company Solvency and Performance Using a Statistical Model: A Comparative UK Based Study," *Accounting and Business Research*, 13, 295–307.

——— (1984): "Empirical models for the monitoring of UK corporations," *Journal of Banking and Finance*, 8, 199–227.

TIAN, S. AND Y. YU (2017): "Financial ratios and bankruptcy predictions: An international evidence," *International Review of Economics and Finance*, 51, 510–526.

TIAN, S., Y. YU, AND H. GUO (2015): "Variable selection and corporate bankruptcy forecasts," *Journal of Banking and Finance*, 52, 89–100.

TIBSHIRANI, R. (1996): "Regression Shrinkage and Selection Via the Lasso," *Journal of the Royal Statistical Society: Series B (Methodological)*, 58, 267–288.

TIKHONOV, A. N. (1963): "On the Solution of Ill-posed Problems and the Method of Regularization," *Doklady Akademii Nauk*, 151, 501—–504.

TOBBACK, E., J. MOEYERSOMS, M. STANKOVA, AND D. MARTENS (2017): "Bankruptcy prediction for SMEs using relational data," *Decision Support Systems*, 102, 69–81.

TSAI, C. (2009): "Feature selection in bankruptcy prediction," *Knowledge-Based Systems*, 22, 120–127.

WANG, W. AND M. GUEDES (2024): " Firm failure prediction for small and medium-sized enterprises and new ventures," *Review of Managerial Science*.

WERNLI, R. AND A. DIETRICH (2022): "Only the Brave: Improving Self-Rationing Efficiency Among Discouraged Swiss SMEs," *Small Business Economics*, 59, 977–1003.

YOUDEN, W. (1950): "Index for rating diagnostic tests," *Cancer*, 3, 32–35.

ZAVGREN, C. (1983): "The prediction of corporate failure: The state of the art," *Journal of Accounting Literature*, 2.

ZAVGREN, C. V. (1985): "Assessing the Vulnerability to Failure of American Industrial Firms: a Logistic Analysis," *Journal of Business Finance and Accounting*, 12.

ZHAO, J., J. OUENNICHE, AND J. DE SMEDT (2024): "Survey, classification and critical analysis of the literature on corporate bankruptcy and financial distress prediction," *Machine Learning with Applications*, 15.

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

ZOU, H. AND T. HASTIE (2005): "Regularization and Variable Selection Via the Elastic Net," *Journal of the Royal Statistical Society Series B: Statistical Methodology*, 67, 301–320.

Figures and Tables

Table 1:
Definition of Country Groups

This table shows the countries and their groupings covered in this paper. Owing to its size, Turkey constitutes its own entity.

Country Group		Country	ISO
Central Asia	CA	Kazakhstan	KAZ
		Kyrgyz Republic	KGZ
		Mongolia	MNG
		Tajikistan	TJK
		Uzbekistan	UZB
Central and Eastern Europe	CEE	Bulgaria	BGR
		Croatia	HRV
		Czech Republic	CZE
		Estonia	EST
		Hungary	HUN
		Latvia	LVA
		Lithuania	LTU
		Poland	POL
		Romania	ROU
		Slovakia	SVK
		Slovenia	SLN
Eastern Neighbourhood	EN	Armenia	ARM
		Azerbaijan	AZE
		Belarus	BLR
		Georgia	GEO
		Moldova	MDA
		Ukraine	UKR
Southern Neighbourhood	SN	Egypt	EGY
		Jordan	JOR
		Lebanon	LBN
		Morocco	MAR
		Palestine	PSE
		Tunisia	TUN
Western Balkans	WB	Albania	ALB
		Bosnia and Herzegovina	BIH
		Kosovo	XKX
		Montenegro	MNE
		North Macedonia	MKD
		Serbia	SRB

TABLE 2:
NEED FOR LOANS

This table provides evindence on the extent to which firms needing a loan are able to obtain one. Column 1 reports the share of firms that have stated they are in need of a loan. Column 2 reports the share of firms that have stated they applied for a loan. Column 3 reports the share of firms that have stated they had their loan application rejected. Column 4 reports the share of firms that have stated they were discouraged from applying for a loan. Regional results are highlighted in gray.

	NEED [% of firms]	APPLIED [% of firms]	REJECTED [% of firms]	DISCOURAGED [% of firms]
	(1)	(2)	(3)	(4)
CA	36.6	14.3	2.4	22.3
KAZ	32.1	9.7	1.8	22.3
KGZ	27.0	15.3	1.0	11.7
MNG	82.2	44.2	9.8	38.0
TJK	31.1	11.6	1.0	19.4
UZB	38.7	19.0	2.5	19.7
CEE	32.5	19.4	1.1	13.1
BGR	34.6	12.7	0.3	21.9
CZE	28.9	25.9	0.2	3.0
EST	29.9	26.2	3.3	3.7
HRV	29.3	24.8	0.8	4.5
HUN	30.5	23.7	0.3	6.8
LTU	32.9	21.0	3.3	12.0
LVA	32.2	22.8	0.8	9.4
POL	26.7	13.3	0.6	13.4
ROU	48.9	14.3	3.0	34.6
SVK	26.9	13.4	0.5	13.5
SVN	34.2	32.3	1.4	1.9
EN	57.6	21.4	2.6	36.2
ARM	60.6	27.2	0.7	33.4
AZE	31.8	13.5	1.3	18.3
BLR	49.0	30.6	3.7	18.4
GEO	40.6	31.3	3.8	9.3
MDA	54.0	19.0	6.1	35.0
UKR	65.1	15.7	1.7	49.5
SN	29.8	6.7	0.7	23.1
EGY	26.1	4.1	0.6	22.0
JOR	30.8	13.0	2.2	17.8
LBN	53.6	25.7	1.8	27.9
MAR	45.8	15.3	0.8	30.4
PSE	24.1	11.5	1.5	12.6
TUN	59.5	23.8	1.5	35.7
TUR	60.5	23.5	0.9	37.0
WB	37.8	26.8	0.6	10.9
ALB	23.6	18.3	0.0	5.3
BIH	38.7	26.2	1.3	12.6
MKD	36.0	19.9	1.6	16.1
MNE	47.7	24.8	0.1	22.9
SRB	45.2	36.2	0.1	8.9
XKX	29.7	13.5	0.5	16.1
TOTAL	38.2	16.0	1.2	22.2

TABLE 3:
VARIABLE DEFINITIONS - ENTERPRISE SURVEY

This table provides the sources and the definitions of the set of candidate predictors from the Enterprise Survey (ES). Column SOURCE reports either the raw ES question or the existing ES indicator variables used to generate our variables. Column DETAILED DEFINITION describes the generated variable.

	SOURCE	DETAILED DEFINITION
PANEL A: IN-SAMPLE RESPONSE		
Applied	k16	1 if the firm applied for a loan.
Rejected	k20a1	1 if the firm applied for a loan and the loan application was rejected.
PANEL B: OUT-OF-SAMPLE RESPONSE		
Discouraged	k16 & k17	1 if the firm did not apply for a loan because of • complex application procedures or, • unfavorable interest rates or, • stringent collateral requirements or, • insufficient volume and maturity or, • an expected loan rejection.
PANEL C: CANDIDATE SURVEY PREDICTORS		
Acquired External Knowledge	BMh1	1 if the firm acquired external knowledge, e.g. licensed of patents.
Audited	t2	1 if the firm had its financial statements checked by an external auditor.
Building Status: Owned	g6a	1 if the firm owns the building it occupies.
Experienced Losses Due to Theft	i3	1 if the firm experienced losses as a result of physical or online theft.
Experienced Power Outages	in16	1 if the firm experienced power outages.
Exporter	exporter	1 if the firm exports more than 10 percent of sales.
Family Ownership	BMb1	1 if the firm has any family ownership.
Female CEO	gend4	1 if the firm has a female CEO.
Female Owners	gend1	1 if the firm has any female owners.
Firm Age: 0-5 Years	car1	1 if the firm is less than five years old.
Import License Application	j10	1 if the firm has submitted an application to obtain an import license.
Inspected by Tax Officials	reg6	1 if the firm was inspected by tax officials.
Introduced New Product or New Process	t7 & t9	1 if the firm introduced new or improved products or processes.
Leased Fixed Assets	BMk5	1 if the company leased fixed assets.
Legal Status: Other	b1	1 if the firm is neither of the other legal status.
Legal Status: Partnership	b1	1 if the firm is owned and run by more than one person.
Legal Status: Private	b1	1 if the firm has non-traded shares or shares traded privately.
Legal Status: Sole Proprietorship	b1	1 if the firm is owned and run by only one person.
Log(Firm Size)	wk14	Logarithm of the firm's number of all employees.
Main Market: International	e1	1 if the market in which the firm sold its main product is international.
Main Market: Local	e1	1 if the market in which the firm sold its main product is local.
Management at Political Position	BMb5	1 if the firm ever had any management at a political position.
No Formal Registration	infor4	1 if the firm wasn't formally registered when it began operations.
Offering Training to Employees	wk1	1 if the firm had formal training programs for its employees.
Overdraft Facility	k7	1 if the firm has an overdraft facility.
Own Website	t5	1 if the firm has its own website.
Paying for Security	crime1	1 if the firm paid for security services.
Purchased Fixed Assets	k4	1 if the firm purchased any new or used fixed assets.
Quality Certification	t1	1 if the firm has an internationally recognized quality certification.
Sales Expectation Next Year: Decrease	BMd1a	1 if the firm expects its total sales to decrease next year.
Sales Expectation Next Year: Increase	BMd1a	1 if the firm expects its total sales to increase next year.
Secured Government Contract	j6a	1 if the firm secured or attempted to secure a government contract.
Trade Credit	k3f	1 if the firm used trade credit to finance its working capital.
Use of Foreign Technology	e6	1 if the firm used technology licensed from a foreign-owned company.

TABLE 4:
SUMMARY STATISTICS - ENTERPRISE SURVEY

This table reports the mean and the standard deviation of the 34 candidate Enterprise Survey predictors for the applicant firms and for the discouraged firms. Column 5 reports the *t*-statistic to test the difference in means across the two groups.

	APPLICANT		DISCOURAGED		
	MEAN [% of firms]	SD [% of firms]	MEAN [% of firms]	SD [% of firms]	*t*
	(1)	(2)	(3)	(4)	(5)
Acquired External Knowledge	19.2	39.4	8.5	27.8	15.0
Audited	49.0	50.0	36.5	48.1	12.2
Building Status: Owned	72.8	44.5	68.4	46.5	4.7
Experienced Losses Due to Theft	13.1	33.7	6.8	25.2	10.1
Experienced Power Outages	41.2	49.2	30.8	46.2	10.4
Exporter	30.3	46.0	14.2	34.9	18.9
Family Ownership	55.5	49.7	43.0	49.5	12.0
Female CEO	14.7	35.5	16.3	37.0	-2.1
Female Owners	31.8	46.6	27.2	44.5	4.8
Firm Age: 0-5 Years	9.2	28.9	8.9	28.5	0.4
Import License Application	10.4	30.6	5.7	23.2	8.4
Inspected by Tax Officials	47.4	49.9	46.4	49.9	0.9
Introduced New Product or New Process	45.5	49.8	21.7	41.3	24.9
Leased Fixed Assets	31.3	46.4	12.4	33.0	22.3
Legal Status: Other	2.5	15.5	4.2	20.0	-4.6
Legal Status: Partnership	17.0	37.6	18.7	39.0	-2.1
Legal Status: Private	57.5	49.4	47.0	49.9	10.1
Legal Status: Sole Proprietorship	15.6	36.3	24.4	43.0	-10.7
Log(Firm Size)	3.7	1.4	3.1	1.2	22.6
Main Market: International	18.4	38.7	7.7	26.7	15.4
Main Market: Local	34.2	47.4	45.7	49.8	-11.4
Management at Political Position	8.3	27.6	6.6	24.8	3.1
No Formal Registration	3.5	18.3	7.1	25.7	-7.8
Offering Training to Employees	42.2	49.4	23.3	42.3	19.7
Overdraft Facility	53.1	49.9	31.6	46.5	21.3
Own Website	70.5	45.6	51.7	50.0	18.9
Paying for Security	65.5	47.5	48.5	50.0	16.8
Purchased Fixed Assets	59.4	49.1	25.2	43.4	35.5
Quality Certification	34.1	47.4	19.2	39.4	16.3
Sales Expectation Next Year: Decrease	15.6	36.3	20.1	40.1	-5.5
Sales Expectation Next Year: Increase	59.7	49.1	48.4	50.0	10.7
Secured Government Contract	25.4	43.5	15.1	35.8	12.5
Trade Credit	27.6	44.7	26.7	44.2	1.0
Use of Foreign Technology	21.1	40.8	13.5	34.2	9.7

TABLE 5:
COMPARISON OF REJECTION MODEL SPECIFICATIONS
COUNTRY VS MACROECONOMIC VARIABLES

This table compares different model specifications. For each logistic elastic net models with $\alpha \in (0, 1)$, the first specification includes survey, sector and country variables (Columns SSC), and the second specification includes survey, sector and macro variables (Columns SSM). For both specifications, we report the rank and the sign of the selected variables' coefficients ($+$ for positive, $-$ for negative and . for zero, i.e. not selected). Finally, we report the mean deviance in the cross-validated samples and the area under curve (AUC) for the full model, i.e. no penalty $\lambda = 0$, and the penalized model, i.e. with penalty $\lambda > 0$. The penalty term λ is selected via a leave-one-out cross validation.

	PENALIZED LOGISTIC REGRESSION WITH LEAVE-ONE-OUT CROSS-VALIDATION																			
	SSC										SSM									
	RIDGE		$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.75$		LASSO		RIDGE		$\alpha = 0.25$		$\alpha = 0.50$		$\alpha = 0.75$		LASSO	
	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$	#	$\pm$
	(1)		(2)		(3)		(4)		(5)		(6)		(7)		(8)		(9)		(10)	
Overdraft Facility	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-	1	-
Purchased Fixed Assets	2	-	2	-	2	-	2	-	2	-	3	-	2	-	2	-	2	-	2	-
Sales Expectation Next Year: Decrease	5	+	5	+	4	+	3	+	3	+	6	+	6	+	5	+	5	+	5	+
Building Status: Owned	4	-	4	-	3	-	4	-	4	-	4	-	4	-	4	-	4	-	4	-
Import License Application	3	-	3	-	5	-	5	-	5	-	2	-	3	-	3	-	3	-	3	-
Firm Age: 0-5 Years	6	+	6	+	6	+	6	+	6	+	5	+	5	+	6	+	6	+	6	+
Log(Firm Size)	8	-	8	-	7	-	7	-	7	-	8	-	8	-	8	-	8	-	8	-
Experienced Losses Due to Theft	7	+	7	+	8	+	8	+	8	+	7	+	7	+	7	+	7	+	7	+
Quality Certification	9	-	9	-	9	-	9	-	9	-	9	-	9	-	9	-	9	-	9	-
Leased Fixed Assets	10	-	10	-	10	-	10	-	10	-	10	-	11	-	10	-	10	-	10	-
Secured Government Contract	17	+	12	+	12	+	12	+	11	+	14	+	13	+	12	+	12	+	11	+
Own Website	11	-	11	-	11	-	11	-	12	-	11	-	10	-	11	-	11	-	13	-
Acquired External Knowledge	30	+	.	.	.	.	.	.	.	.	23	+	.	.	.	.	.	.	.	.
Audited	14	-	16	-	.	.	.	.	.	.	20	-	17	-	17	-	17	-	17	-
Experienced Power Outages	25	+	.	.	.	.	.	.	.	.	30	+	.	.	.	.	.	.	.	.
Exporter	29	-	.	.	.	.	.	.	.	.	26	-	.	.	.	.	.	.	.	.
Family Ownership	13	-	15	-	14	-	.	.	.	.	25	-	.	.	.	.	.	.	.	.
Female CEO	19	-	.	.	.	.	.	.	.	.	19	-	.	.	.	.	.	.	.	.
Female Owners	15	+	17	+	.	.	.	.	.	.	16	+	16	+	16	+	16	+	16	+
Inspected by Tax Officials	27	+	.	.	.	.	.	.	.	.	33	+	.	.	.	.	.	.	.	.
Introduced New Product or New Process	28	+	.	.	.	.	.	.	.	.	22	+	.	.	.	.	.	.	.	.
Legal Status: Other	22	-	.	.	.	.	.	.	.	.	21	+	12	+	14	+	14	+	14	+
Legal Status: Partnership	21	-	.	.	.	.	.	.	.	.	28	-	.	.	.	.	.	.	.	.
Legal Status: Private	20	-	.	.	.	.	.	.	.	.	15	-	.	.	.	.	.	.	.	.
Legal Status: Sole Proprietorship	18	-	.	.	.	.	.	.	.	.	13	-	15	-	15	-	15	-	15	-
Main Market: International	23	-	.	.	.	.	.	.	.	.	24	-	19	-	.	.	.	.	.	.
Main Market: Local	33	+	.	.	.	.	.	.	.	.	34	+	.	.	.	.	.	.	.	.
Management at Political Position	26	+	.	.	.	.	.	.	.	.	17	+	.	.	.	.	.	.	.	.
No Formal Registration	32	+	.	.	.	.	.	.	.	.	32	+	.	.	.	.	.	.	.	.
Offering Training to Employees	34	+	.	.	.	.	.	.	.	.	27	+	.	.	.	.	.	.	.	.
Paying for Security	12	-	13	-	13	-	13	-	.	.	18	-	18	-	18	-	18	-	.	.
Sales Expectation Next Year: Increase	31	-	.	.	.	.	.	.	.	.	31	-	.	.	.	.	.	.	.	.
Trade Credit	16	+	14	+	.	.	.	.	.	.	12	+	14	+	13	+	13	+	12	+
Use of Foreign Technology	24	-	.	.	.	.	.	.	.	.	29	+	.	.	.	.	.	.	.	.
NUMBER OF SELECTED VARIABLES: TOTAL	75/75		48/75		39/75		37/75		36/75		47/47		28/47		26/47		26/47		25/47	
NUMBER OF SELECTED VARIABLES: SURVEY	34/34		17/34		14/34		13/34		12/34		34/34		19/34		18/34		18/34		17/34	
NUMBER OF SELECTED VARIABLES: SECTOR	6/6		4/6		3/6		3/6		3/6		6/6		5/6		4/6		4/6		4/6	
NUMBER OF SELECTED VARIABLES: COUNTRY/MACRO	35/35		27/35		22/35		21/35		21/35		7/7		4/7		4/7		4/7		4/7	
DEVIANCE: UNPENALIZED MODEL $\lambda = 0$	0.4942		0.4942		0.4942		0.4942		0.4942		0.4870		0.4870		0.4870		0.4870		0.4870	
DEVIANCE: PENALIZED MODEL $\lambda > 0$	0.4755		0.4728		0.4722		0.4717		0.4715		0.4835		0.4808		0.4801		0.4798		0.4797	
AUC: UNPENALIZED MODEL $\lambda = 0$	0.76		0.76		0.76		0.76		0.76		0.75		0.75		0.75		0.75		0.75	
AUC: PENALIZED MODEL $\lambda > 0$	0.82		0.81		0.81		0.81		0.81		0.79		0.78		0.78		0.78		0.78	

TABLE 6: COMPARISON OF REJECTION MODEL SPECIFICATIONS WITH OR WITHOUT REGIONAL INTERACTIONS

For each logistic Lasso model, Columns 1-2 report the first specification with survey, sector and country variables, and Columns 3-12 report the second specification with the regional interactions of survey, sector and country variables. For both specifications, we report the rank and the sign of the selected variables' coefficients ($+$ for positive, $-$ for negative and . for zero, i.e. not selected). Finally, we report the mean deviance in the cross-validated samples and the area under curve (AUC) for the full model, i.e. no penalty $\lambda = 0$, and the penalized model, i.e. with penalty $\lambda > 0$. The penalty term λ is selected via a leave-one-out cross validation.

LASSO	SSC		RSSC									
			CA		CEE		EN		SN		WB	
	#	±	#	±	#	±	#	±	#	±	#	±
	(1)		(2)		(3)		(4)		(5)		(6)	
Overdraft Facility	1	-	10	-	1	-	2	-	2	-	2	-
Purchased Fixed Assets	2	-	7	-	3	-	10	-	10	-	.	.
Sales Expectation Next Year: Decrease	3	+	1	+	.	.	4	+	4	+	10	+
Building Status: Owned	4	-	13	-	6	-	13	-	18	-	5	-
Import License Application	5	-	2	-	2	-	15	-	13	-	.	.
Firm Age: 0-5 Years	6	+	3	+	4	+	3	+	16	-	9	-
Log(Firm Size)	7	-	17	-	14	-	5	-	.	.	.	.
Experienced Losses Due to Theft	8	+	8	+	8	+	8	+	1	-	.	.
Quality Certification	9	-	4	-	.	.	.	.	5	-	.	.
Leased Fixed Assets	10	-	5	-	.	.	12	-	8	-	.	.
Secured Government Contract	11	+	9	+	.	.	.	.	.	.	7	-
Own Website	12	-	.	.	13	-	.	.	.	.	6	-
Acquired External Knowledge	.	.	11	-	.	.	7	+	.	.	.	.
Audited	.	.	.	.	10	-	.	.	6	-	.	.
Experienced Power Outages	.	.	6	+	.	.	19	+	.	.	4	-
Exporter	.	.	.	.	.	.	.	.	17	-	3	-
Family Ownership	.	.	.	.	.	.	.	.	.	.	8	-
Female CEO	.	.	.	.	.	.	17	-	.	.	.	.
Female Owners	.	.	.	.	.	.	16	+	15	+	.	.
Inspected by Tax Officials	.	.	.	.	.	.	.	.	.	.	.	.
Introduced New Product or New Process	.	.	.	.	.	.	.	.	.	.	.	.
Legal Status: Other	.	.	.	.	.	.	.	.	.	.	.	.
Legal Status: Partnership	.	.	.	.	5	+	.	.	.	.	1	-
Legal Status: Private	.	.	.	.	.	.	11	+	9	-	.	.
Legal Status: Sole Proprietorship	.	.	.	.	12	-	6	-	7	+	.	.
Main Market: International	.	.	.	.	7	-	.	.	3	+	.	.
Main Market: Local	.	.	19	+	.	.	.	.	.	.	.	.
Management at Political Position	.	.	18	+	.	.	.	.	14	-	.	.
No Formal Registration	.	.	15	-	.	.	.	.	12	+	11	-
Offering Training to Employees	.	.	.	.	.	.	14	+	.	.	.	.
Paying for Security	.	.	16	+	9	-	18	-	.	.	.	.
Sales Expectation Next Year: Increase	.	.	12	-	.	.	.	.	11	+	.	.
Trade Credit	.	.	14	-	11	+	9	+	.	.	.	.
Use of Foreign Technology	.	.	.	.	.	.	1	-	.	.	.	.
NUMBER OF SELECTED VARIABLES: TOTAL	36/75		26/45		22/51		26/46		25/46		14/46	
NUMBER OF SELECTED VARIABLES: SURVEY	12/34		19/34		14/34		19/34		18/34		11/34	
NUMBER OF SELECTED VARIABLES: SECTOR	3/6		4/6		2/6		3/6		4/6		2/6	
NUMBER OF SELECTED VARIABLES: COUNTRY	21/35		3/5		6/11		4/6		3/6		1/6	
DEVIANCE: UNPENALIZED MODEL $\lambda = 0$	0.49						0.65					
DEVIANCE: PENALIZED MODEL $\lambda > 0$	0.47						0.48					
AUC: UNPENALIZED MODEL $\lambda = 0$	0.76						0.73					
AUC: PENALIZED MODEL $\lambda > 0$	0.81						0.86					

TABLE 7: CREDIT ALLOCATION VIA MINIMIZING MISCLASSIFICATION COST

This table reports the credit allocation as a result of misclassification cost minimization. The misclassification cost is defined as:

$$\text{MISCLASSIFICATION COST} = \text{RELATIVE COST} \cdot \text{FALSE NEGATIVE RATE} \cdot P(\text{REJECTED}) + \text{FALSE POSITIVE RATE} \cdot P(\text{ACCEPTED})$$

Column 1 is the relative cost, i.e. the cost of lending to a rejected firm relative to the opportunity cost of not lending to an accepted firm. Column 2 and 3 report the threshold rejection probability with the minimum misclassification cost for the given relative cost. Columns 4-7 report the false negative rate and the false positive rate resulting from a credit allocation with the selected threshold rejection probability. Note that the positive case in the classification after the credit allocation is being rejected. Columns 8-11 report the share of rejections in the predicted outcomes in-sample and out-of-sample, i.e. for discouraged firms. SSC (RSSC) column highlights the specification with (regional interactions of) survey, sector and country variables.

RELATIVE COST	REJECTION THRESHOLD		FALSE NEGATIVE RATE		FALSE POSITIVE RATE		PREDICTED REJECTIONS			
							IN-SAMPLE		OUT-OF-SAMPLE	
	SSC	RSSC	SSC	RSSC	SSC	RSSC	SSC	RSSC	SSC	RSSC
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)
0	59.4%	60.0%	100.0%	99.3%	0.0%	0.0%	0.0%	0.1%	0.1%	0.3%
1	33.5%	32.0%	90.3%	83.8%	0.7%	0.8%	1.4%	2.0%	6.1%	8.2%
2	31.8%	30.4%	88.1%	80.9%	0.9%	1.0%	1.7%	2.4%	7.5%	9.7%
3	22.9%	20.1%	74.4%	58.1%	3.5%	5.5%	5.2%	8.3%	18.2%	25.1%
4	16.6%	16.5%	56.3%	46.6%	9.3%	9.2%	12.0%	12.6%	32.7%	34.1%
5	16.2%	15.5%	54.9%	42.6%	9.9%	10.6%	12.6%	14.3%	33.6%	36.9%
6	13.3%	15.5%	45.8%	42.6%	14.1%	10.6%	17.2%	14.3%	45.0%	36.9%
7	10.3%	13.6%	31.8%	35.7%	21.8%	14.3%	25.4%	18.2%	56.6%	43.8%
8	9.4%	10.3%	27.1%	23.8%	24.8%	21.9%	28.5%	26.1%	60.5%	56.3%
9	9.4%	10.3%	27.1%	23.8%	24.8%	21.9%	28.5%	26.1%	60.5%	56.3%
10	8.9%	9.4%	24.9%	19.5%	26.5%	25.2%	30.3%	29.5%	62.5%	60.2%
11	7.6%	9.4%	19.1%	19.5%	31.8%	25.2%	35.6%	29.5%	67.4%	60.2%
12	7.6%	9.4%	19.1%	19.5%	31.8%	25.2%	35.6%	29.5%	67.4%	60.2%
13	7.6%	9.4%	19.1%	19.5%	31.8%	25.2%	35.6%	29.5%	67.4%	60.2%
14	7.6%	8.2%	19.1%	16.2%	31.8%	28.9%	35.6%	33.1%	67.4%	64.3%
15	7.6%	8.2%	19.1%	16.2%	31.8%	28.9%	35.6%	33.1%	67.4%	64.3%

TABLE 8: CREDIT GAP ESTIMATES PROPORTIONAL CREDIT ALLOCATION

This table reports the credit gaps resulting from an allocation where each firm gets credit in proportion to their approval probabilities, i.e. every firm gets rationed. Columns 1-4 report the baseline credit gap in percent of GDP and in million US dollars, respectively. Columns 5-8 report the SME credit gap in percent of GDP and in million US dollars, respectively. Columns 9-10 show the share of the SME credit gap in the total credit gap. SSC (RSSC) columns highlight the specifications with (regional interactions of) survey, sector and country variables. Regional results are highlighted in gray.

	CREDIT GAP				SME CREDIT GAP					
	% GDP		MILLION USD		% GDP		MILLION USD		% BASELINE	
	SSC	RSSC	SSC	RSSC	SSC	RSSC	SSC	RSSC	SSC	RSSC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
CA	4.7	4.7	12,204	12,147	3.6	3.6	9,307	9,334	76	77
KAZ	4.4	4.4	7,889	7,877	3.5	3.5	6,246	6,279	79	80
KGZ	2.1	2.0	175	166	1.3	1.2	110	103	63	62
MNG	3.8	3.8	498	502	2.7	2.7	351	356	71	71
TJK	3.6	3.6	280	281	2.4	2.4	190	187	67	67
UZB	6.4	6.3	3,361	3,320	4.6	4.6	2,414	2,406	72	72
CEE	4.6	4.5	76,380	73,433	3.0	3.0	49,849	49,360	65	67
BGR	12.3	12.4	8,182	8,248	9.8	9.9	6,539	6,605	80	80
CZE	0.2	0.2	578	578	0.2	0.2	502	500	87	86
EST	0.5	0.5	151	149	0.4	0.4	120	118	79	79
HRV	0.4	0.4	252	254	0.3	0.3	209	210	83	83
HUN	1.1	1.1	1,779	1,777	0.7	0.7	1,162	1,170	65	66
LTU	2.5	2.6	1,368	1,423	2.4	2.5	1,283	1,337	94	94
LVA	1.7	1.6	590	568	1.3	1.3	449	435	76	76
POL	4.7	4.6	27,354	27,230	4.3	4.3	25,242	25,225	92	93
ROU	8.1	7.7	19,600	18,483	4.5	4.3	10,906	10,420	56	56
SVK	15.5	13.8	16,411	14,604	3.1	3.1	3,321	3,323	20	22
SVN	0.2	0.2	115	118	0.2	0.2	115	118	100	100
EN	9.1	9.1	25,426	25,494	5.0	5.1	14,018	14,130	55	55
ARM	3.4	3.4	418	423	1.5	1.6	191	194	46	46
AZE	1.3	1.3	611	602	1.1	1.1	507	497	83	83
BLR	3.6	3.6	2,136	2,164	1.1	1.1	658	682	31	32
GEO	0.9	0.9	152	158	0.9	0.9	152	158	100	100
MDA	6.0	6.1	679	683	2.1	2.0	231	230	34	34
UKR	16.4	16.4	21,429	21,463	9.4	9.4	12,279	12,370	57	58
SN	19.8	19.6	108,196	107,160	13.6	13.5	74,567	73,806	69	69
EGY	16.2	15.9	42,582	41,898	13.0	12.8	34,205	33,738	80	81
JOR	43.5	42.4	18,722	18,213	19.9	19.0	8,559	8,179	46	45
LBN	24.3	23.4	13,324	12,855	20.9	20.1	11,500	11,011	86	86
MAR	21.2	22.0	26,966	27,961	12.9	13.5	16,488	17,240	61	62
PSE	11.1	10.1	1,799	1,641	9.7	9.0	1,584	1,460	88	89
TUN	11.3	10.8	4,802	4,592	5.2	5.1	2,230	2,178	46	47
TUR	12.4	12.5	96,910	97,700	10.0	10.1	77,855	78,728	80	81
WB	2.6	2.6	2,901	2,911	1.6	1.6	1,794	1,815	62	62
ALB	0.8	0.9	127	129	0.7	0.7	102	104	81	81
BIH	4.4	4.3	893	867	2.3	2.3	471	460	53	53
MKD	2.3	2.4	298	301	1.9	2.0	244	248	82	82
MNE	1.3	1.3	70	69	1.3	1.3	70	69	100	100
SRB	1.7	1.6	838	834	0.7	0.8	378	382	45	46
XKX	8.6	9.0	675	710	6.7	7.0	528	552	78	78
TOTAL	8.9	8.8	322,016	318,844	6.3	6.3	227,390	227,173	71	71

TABLE 9:
CREDIT GAP ESTIMATES
THRESHOLD-BASED CREDIT ALLOCATION

This table reports the credit gaps resulting from an allocation based on the threshold set at a relative cost of 11 for lending to a rejected firm following subsection 4.3. Columns 1-4 report the total credit gap in percent of GDP and in million US dollars, respectively. Columns 5-6 report these credit gap estimates as percentage of the credit gap estimate resulting from matching the sample rejections. Finally, Columns 7-8 report the percentage point differences from the matched credit gap estimates. SSC (RSSC) columns highlight the specifications with (regional interactions of) survey, sector and country variables. Regional results are highlighted in gray.

	CREDIT GAP							
	% GDP		MILLION USD		% BASELINE		± PP. FROM BASELINE	
	SSC	RSSC	SSC	RSSC	SSC	RSSC	SSC	RSSC
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
CA	1.2	1.2	3,196	3,156	33	33	-3.4	-3.4
KAZ	1.0	0.8	1,810	1,374	25	5	-3.4	-3.6
KGZ	1.1	0.7	89	57	38	29	-1.0	-1.3
MNG	1.8	2.4	236	312	47	50	-2.0	-1.4
TJK	1.9	2.4	147	189	43	60	-1.7	-1.2
UZB	1.7	2.3	914	1,224	42	57	-4.7	-4.0
CEE	2.2	2.1	36,646	34,001	57	63	-2.4	-2.4
BGR	7.9	7.5	5,242	4,983	70	71	-4.4	-4.9
CZE	0.1	0.2	364	383	78	79	-0.1	-0.1
EST	0.3	0.3	94	87	67	64	-0.2	-0.2
HRV	0.4	0.4	239	263	81	83	0.0	0.0
HUN	1.0	1.0	1,625	1,543	61	63	-0.1	-0.1
LTU	0.0	0.1	0	48	0	0	0.0	-2.6
LVA	0.6	0.5	208	156	44	83	-1.1	-1.2
POL	2.3	2.4	13,507	14,180	86	90	-2.4	-2.2
ROU	3.3	2.6	7,931	6,318	26	33	-4.8	-5.0
SVK	7.0	5.6	7,361	5,920	25	23	-8.6	-8.2
SVN	0.1	0.2	77	121	100	100	-0.1	0.0
EN	3.7	4.3	10,345	11,931	30	36	-5.4	-4.9
ARM	2.9	3.2	366	395	36	42	-0.4	-0.2
AZE	0.6	0.7	292	321	70	72	-0.7	-0.6
BLR	3.1	3.3	1,854	1,970	20	23	-0.5	-0.3
GEO	0.0	0.2	0	36	0	100	-0.0	-0.7
MDA	3.1	4.3	344	482	5	9	-3.0	-1.8
UKR	5.7	6.7	7,489	8,728	32	38	-10.6	-9.7
SN	6.5	9.6	35,594	52,501	55	62	-13.3	-10.0
EGY	6.3	6.7	16,613	17,565	55	62	-9.9	-9.2
JOR	0.8	6.4	348	2,763	82	56	-42.7	-35.9
LBN	11.3	12.6	6,211	6,893	74	73	-13.0	-10.9
MAR	7.9	18.1	10,112	23,075	46	62	-13.2	-3.8
PSE	4.2	2.5	689	406	84	77	-6.8	-7.6
TUN	3.8	4.2	1,621	1,801	33	42	-7.5	-6.5
TUR	11.5	12.6	89,353	97,984	79	80	-1.0	0.0
WB	2.3	2.3	2,607	2,524	59	62	-0.3	-0.3
ALB	0.8	0.9	128	130	81	81	0.0	0.0
BIH	4.3	3.4	864	682	52	55	-0.1	-0.9
MKD	1.5	1.6	185	198	70	77	-0.9	-0.8
MNE	1.2	1.0	65	55	100	100	-0.1	-0.3
SRB	1.6	1.6	796	828	42	46	-0.1	0.0
XKX	7.2	8.0	568	632	82	80	-1.4	-1.0
TOTAL	4.9	5.6	177,742	202,096	66	69	-4.0	-3.2

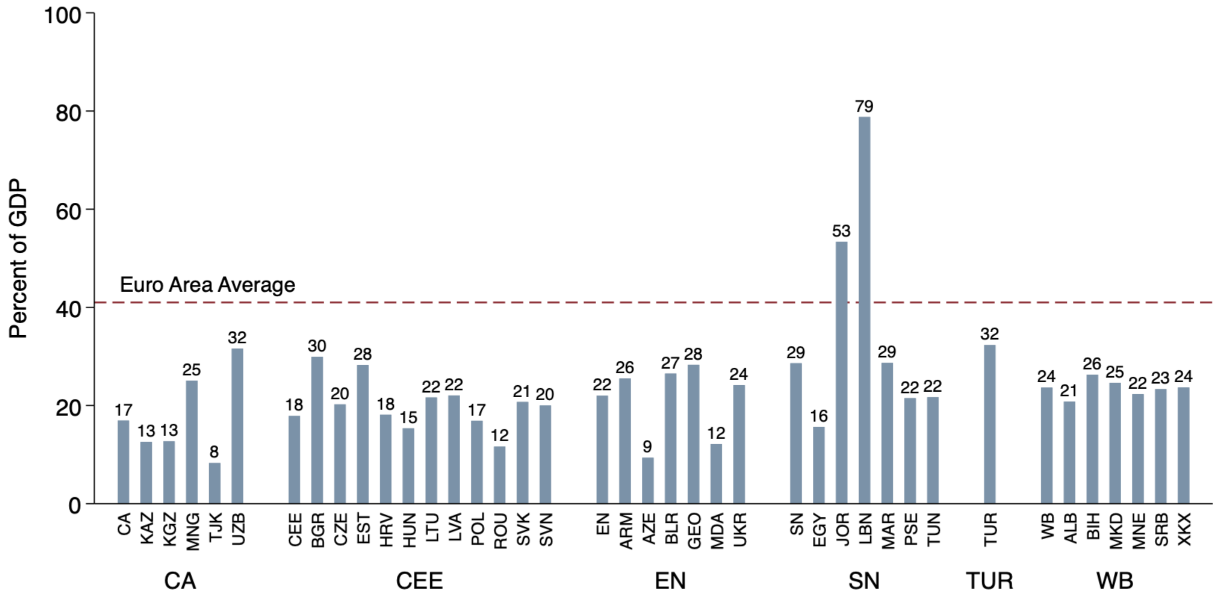


FIGURE 1: CREDIT TO NON-FINANCIAL CORPORATIONS

This figure plots the credit to non-financial corporations (NFC) relative to GDP for each country and region in our sample. The data primarily come from the International Monetary Fund's (IMF) the Financial Soundness Indicators (FSI) and the European Central Bank (ECB); when these resources are not available for a country we resort to the IMF's the Financial Access Survey (FAS) and the local central banks. The stock of NFC credit is adjusted by the share of value added in the industrial and services sectors, which is obtained from the World Bank (WB). The red line indicates the Euro area average of NFC credit to GDP.

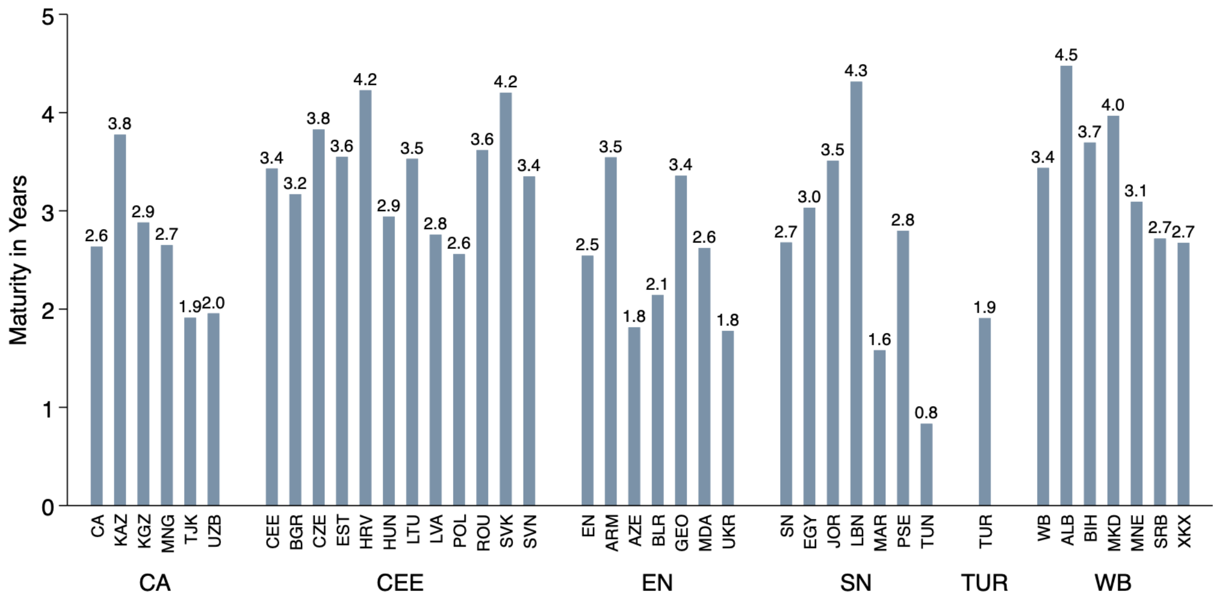


FIGURE 2: AVERAGE ORIGINAL MATURITY OF LOANS

This figure shows the average maturity in years of the last outstanding loan for firms in the sample. The data come from firms' responses to Q.BMk10 in the 2018-2020 wave of the Enterprise Surveys.

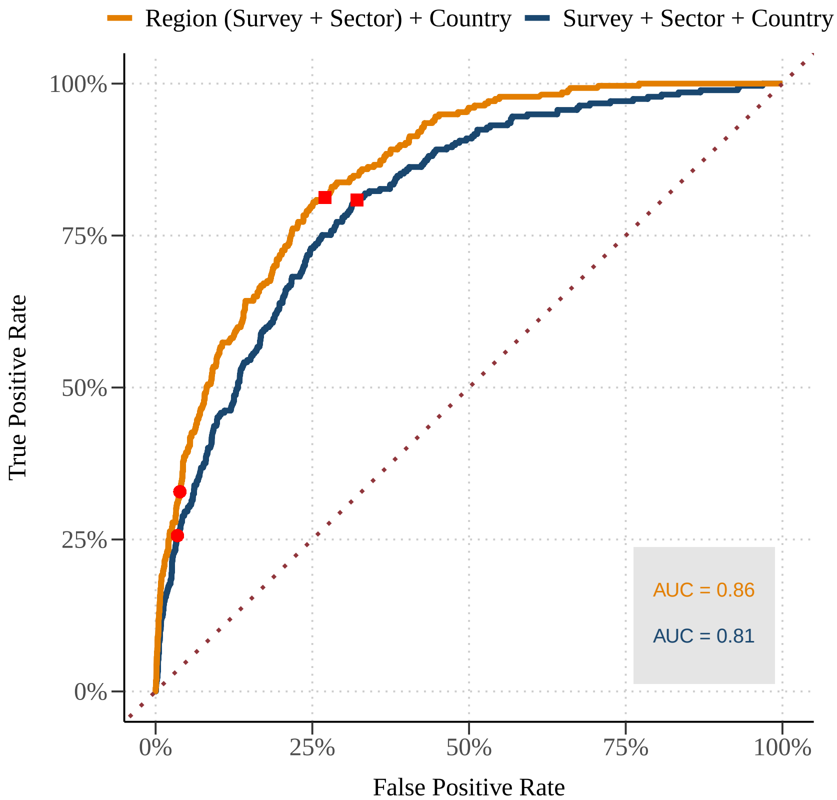
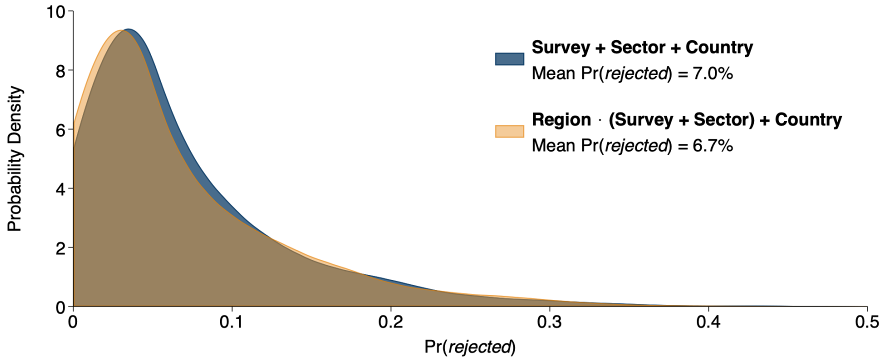
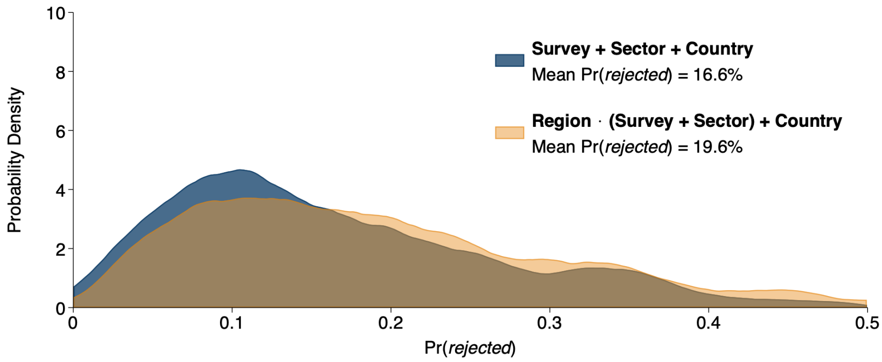


FIGURE 3:
REJECTION PREDICTION PERFORMANCE

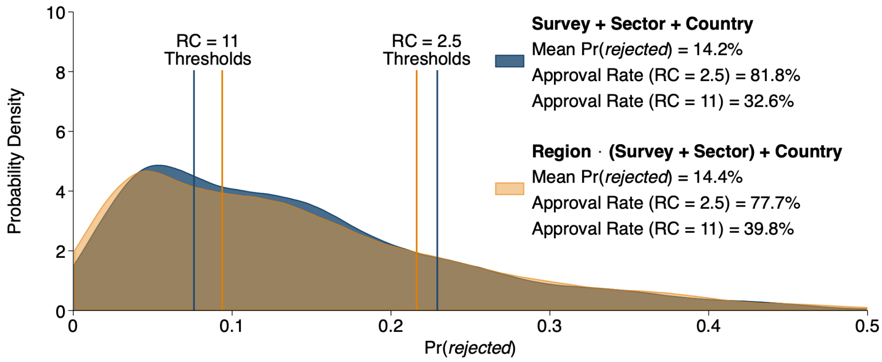
This figure plots the receiver operating characteristic (ROC) curves the model specification with survey, sector and country variables (blue line) and with regional interactions of survey, sector and country variables (orange line), hence illustrating the performance of the rejection predictions from each model specification relative to the actual rejections in the data. The dotted red line is the 45 degree line. The y-axis is the true positive rate (TPR), i.e. the rate of classifying actually rejected firms as rejected. The x-axis is the false positive rate (FPR), i.e. the rate of classifying actually approved firms as rejected. On the ROC curve, we highlight the thresholds set at a relative cost of 2.5 (●) and 11 (■).



(A) APPROVED APPLICANTS



(B) REJECTED APPLICANTS



(C) DISCOURAGED FIRMS

FIGURE 4:
PREDICTED REJECTION PROBABILITIES AND CREDIT ALLOCATION

This figure shows the distributions of firms' rejection probabilities in the (Survey + Sector + Country) model (blue) and the (Region · (Survey + Sector) + Country) model (orange). Panels A and B present the rejection probabilities in-sample for firms whose loan applications were approved and rejected, respectively. Panel C presents the rejection probabilities out-of-sample for firms that were discouraged from applying for a loan. In each panel, we report the mean rejection probability. Furthermore, Panel C shows how credit is allocated to discouraged firms; a discouraged firm with an estimated rejection probability below a rejection threshold gets credit. An allocation based on a proportional mechanism is equivalent to one based on a threshold mechanism, where the rejection threshold is set at a relative cost of 2.5. We further highlight the rejection threshold set at a relative cost of 11 as a more conservative alternative.

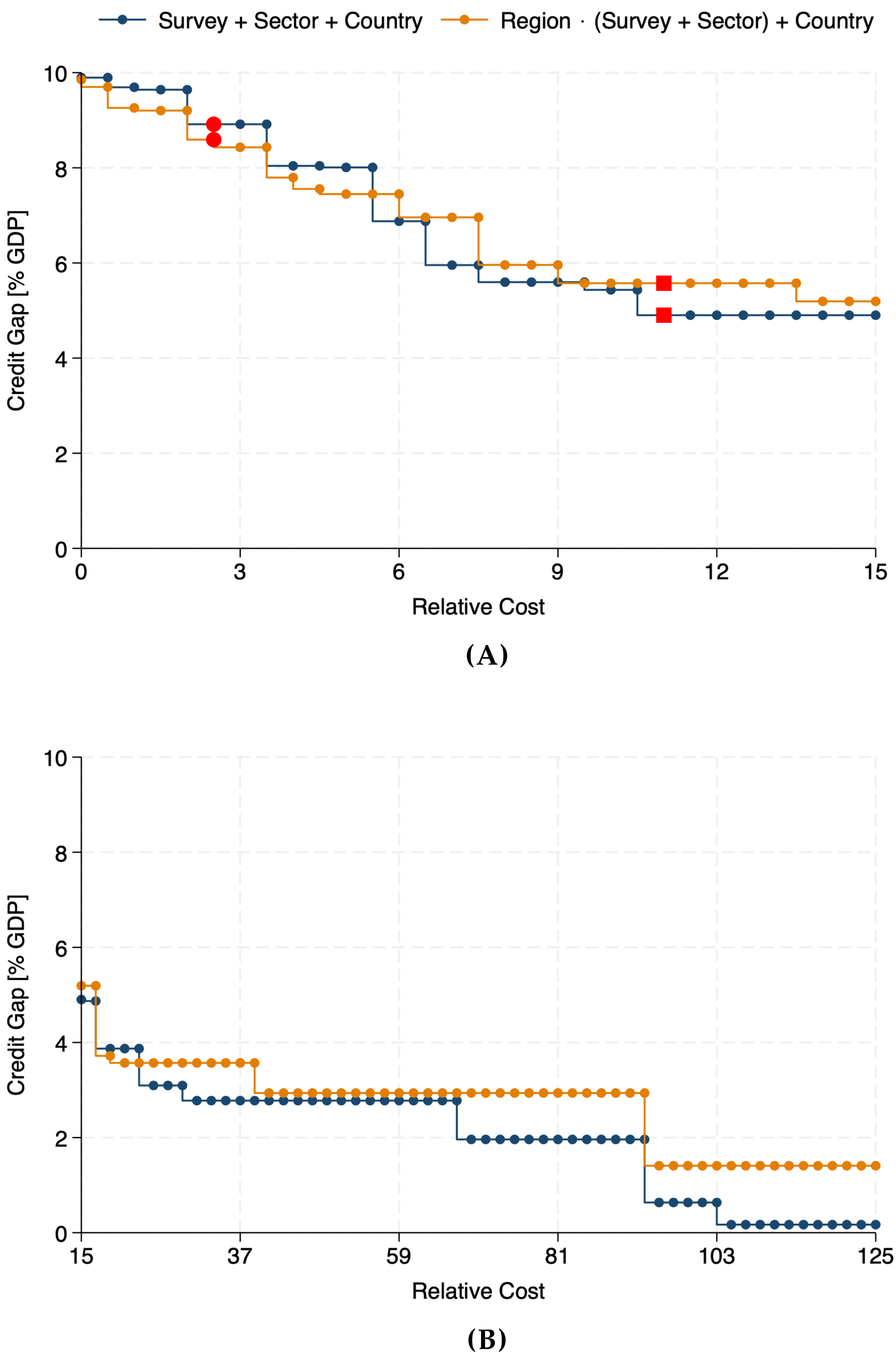


FIGURE 5:
CREDIT GAPS BY RELATIVE COST

This figure plots the credit gap estimates using the (Survey + Sector + Country) model (blue) and the (Region · (Survey + Sector) + Country) model (orange) as a function of the relative cost of lending to a rejected firm. We highlight the relative cost of 2.5 (●) and 11 (■).