BANCA D'ITALIA

EUROSISTEMA

Temi di discussione

(Working Papers)

Machine learning in the service of policy targeting: the case of public credit guarantees

by Monica Andini, Michela Boldrini, Emanuele Ciani, Guido de Blasio, Alessio D'Ignazio and Andrea Paladini

February 2019

Number 1206

BANCA D'ITALIA
EUROSISTEMA

Temi di discussione

(Working Papers)

Machine learning in the service of policy targeting: the case of public credit guarantees

by Monica Andini, Michela Boldrini, Emanuele Ciani,
Guido de Blasio, Alessio D'Ignazio and Andrea Paladini

Number 1206 - February 2019

The papers published in the Temi di discussione *series describe preliminary results and are made available to the public to encourage discussion and elicit comments.*

The views expressed in the articles are those of the authors and do not involve the responsibility of the Bank.

ISSN 1594-7939 (print)
ISSN 2281-3950 (online)

Printed by the Printing and Publishing Division of the Bank of Italy

MACHINE LEARNING IN THE SERVICE OF POLICY TARGETING: THE CASE OF PUBLIC CREDIT GUARANTEES

by Monica Andini[*], Michela Boldrini[**], Emanuele Ciani[*], Guido de Blasio[*], Alessio D'Ignazio[*], Andrea Paladini[***]

Abstract

We use Machine Learning (ML) predictive tools to propose a policy-assignment rule designed to increase the effectiveness of public guarantee programs. This rule can be used as a benchmark to improve targeting in order to reach the stated policy goals. Public guarantee schemes should target firms that are both financially constrained and creditworthy, but they often employ naïve assignment rules (mostly based only on the probability of default) that may lead to an inefficient allocation of resources. Examining the case of Italy's Guarantee Fund, we suggest a benchmark ML-based assignment rule, trained and tested on credit register data. Compared with the current eligibility criteria, the ML-based benchmark leads to a significant improvement in the effectiveness of the Fund in gaining credit access to firms. We discuss the problems in estimating and using these algorithms for the actual implementation of public policies, such as transparency and omitted payoffs.

JEL Classification: C5, H81.
Keywords: machine learning, program evaluation, loan guarantees.

Contents

[*] Bank of Italy, Structural Economic Analysis Directorate.
[**] University of Bologna, Department of Economics.
[***] University of Rome "La Sapienza", Department of Mathematics.

1. Introduction[1]

Machine Learning (ML) tools (Hastie et al., 2009; Varian, 2014) are increasingly used to address prediction problems in applied econometrics.[2] In some instances researchers have a purely forecasting purpose, but the data are quite large or non-conventional. For example, Glaeser et al. (2017) forecast economic activity at very detailed levels of geographic stratification using Yelp data. ML algorithms are also useful for causal inference tasks that have an embedded prediction problem, for instance when we assume that the counterfactual outcomes depend on a very large set of covariates (Belloni et al., 2014). Finally, ML techniques can be used to assist decision makers, by providing them with a decision rule that summarizes the available evidence in order to predict which choice is more likely to serve the purpose. This task is what Kleinberg et al. (2015) define as "prediction policy problem". Chalfin et al. (2016) estimate an algorithm to help hire teachers that are more likely to have higher value added. Kleinberg et al. (2018) study how to use these rules to assist judges in deciding whether or not to grant bail, by exploiting observable information about the accused. When the decision concerns policy targeting, ML methods can be employed ex-ante to identify, among the potential beneficiaries, those who will likely behave in such a way as to ensure the effectiveness of the intervention. Andini et al. (2018), for instance, study the use of ML for targeting a tax bonus intended to spur consumption.

In this paper we focus on the "prediction policy problem" of assigning public credit guarantees to firms. Public guarantee schemes aim to support firms' access to bank credit by providing publicly funded collateral. They typically target small and medium-sized enterprises (SMEs), which are the kind of firms most likely to suffer from credit constraints. These programs, which are widespread in both developed and developing countries, experienced a dramatic surge in popularity in the aftermath of the global financial crisis (Beck et al., 2008). The literature has highlighted that these schemes often fail to reach firms that are actually credit constrained (see, for instance, Zia, 2008). If the guarantee is provided to firms that are not credit constrained the additionality of the program will languish as these firms would have obtained funding anyway. One of the reasons for this misallocation is that credit rationing is difficult to gauge, while firms' creditworthiness is more easily assessed by means of balance sheet variables. As a result, the eligibility condition usually winds down into naïve rules that pinpoint financially sound borrowers, without considering indicators for credit constraints (OECD, 2013). Our

[1] This paper was partly written while Michela Boldrini and Andrea Paladini were interns at the Structural Economic Analysis Directorate of the Bank of Italy, and while Alessio D'Ignazio was visiting the Quantitative and Applied Spatial Economic Research Laboratory (QASER) at University College London. For useful comments and suggestions, we would like to thank Federico Cingano, Giuseppe Grande, Francesca Medda, Fabio Parlapiano, Marco Percoco, Paolo Sestito, Enrico Sette, Luigi Federico Signorini and the participants at the Bank of Italy workshop *Financial factors in the context of economic recovery* (Rome, March 2018), the Bank of Italy workshop on *Big Data & Machine Learning* (Rome, June 2018), the *XXX Annual Conference of the Italian Society of Public Economics* (Padua, September 2018), the conference on *Counterfactual Methods for Policy Impact Evaluation 2018* (Berlin, September 2018), the international conference on *Entrepreneurship and Economic Development: Assessing the Effectiveness of Public Policies* (Bari, October 2018). The views expressed in this paper are those of the authors and do not necessarily correspond to those of the institutions they are affiliated with.

[2] In this paper we use "prediction" and "forecast" interchangeably.

exercise aims at suggesting a benchmark assignment mechanism, based on ML algorithms, that explicitly accounts for both credit constraints and creditworthiness. The nature of this task is essentially a forecasting one. As underscored by Mullainathan and Spiess (2017), these prediction policy problems are the ones for which the ML machinery is extremely well equipped.

We compare our ML-based assignment mechanism against the rule originally put in place. The advantages of the ML tool we propose can be shown by comparing its performance to that of the current allocation rule adopted by the Italian Guarantee Fund (GF) to select firms eligible for public support when accessing credit. First introduced in 2000, the Fund became especially popular with the unfolding of the financial crisis, as the total amount of guarantees that were granted rose from about €1.2 billion in 2008 to €11.6 in 2016. It is the single most important Public Guarantee program in the country.

It is worth noting that the Italian Guarantee Fund was reformed in December 2017. Not unlike the approach we suggest, the reform was mainly designed to improve the screening of creditworthy firms and to increase support to firms that are more likely to be credit constrained. The first objective was addressed by the adoption of a new rating model. The second by granting a higher coverage to riskier (but still creditworthy) firms, assuming that they are more likely to be credit constrained. The reform is still not operative at the time of writing. It is expected to become operative by end 2018. We will leave the comparison between our ML benchmark targeting and the new rules to future research. However, to the extent that the new assignment rule will keep providing guarantees to firms that are not credit constrained, it seems quite possible that the overall effectiveness of the policy could be further improved.

It should also be highlighted that our approach is strictly aimed at improving ex-ante the effectiveness of credit guarantee programs by allocating guarantees to those firms that truly need them. We do not directly target on second round effects, such as investments. We also do not consider general equilibrium effects, which need to be evaluated within a different structural framework.

In the first part of the paper, we work as if we were in the ex-ante situation, in which the policymaker must design the allocation of the guarantee without prior knowledge of the intervention effectiveness. We make use of micro-level data from the credit register (CR), kept at the Bank of Italy, and the Cerved (balance-sheet) dataset, and develop two separate ML prediction models, for credit constraints and creditworthiness, respectively. Hence, all the variables that we use for predicting each status could be potentially available to the GF administration when a firm applies for the guarantee. We try different ML algorithms that are off-the-shelf (Athey, 2018) – LASSO, decision tree, and random forest – and show that the best out-of-sample predictive performances are reached with the latter. The predictions for financially constrained firms are combined with those for creditworthy firms to identify the ML hypothetical beneficiary of the GF. By comparing the GF assignment with the ML assignment, we show that the GF scoring system is biased against firms that are credit constrained. As we explain in detail below, we consider a firm to be credit constrained if the total amount of bank loans granted to that firm does not increase in the six months following a new request for bank credit from that

firm. While this proxy is the best measure of the ease to access bank credit available to us at firm level, it is worth noting that it does not allow us to fully appraise the extent of credit rationing, as the change in the total amount of bank loans reflects both newly issued loans and reimbursement of outstanding loans. In particular, we could overestimate the extent of credit constraints if these reimbursements exceed the amount of new bank loans obtained. A similar error of measurement would materialize if the banks offer new loans to the firm but the firm decides to postpone them or turns to alternative sources of funding.

In the second part of the paper, we substantiate the validity of our approach by looking at the ex-post dimension. As underscored by Athey (2017), ML prediction will not automatically ensure higher program effectiveness because a program might have heterogeneous effects and ML might fail to target those for whom intervention is most beneficial. We provide ex-post empirical evidence to test whether that ML-based assignment mechanism satisfies the aim of increasing the impact of the policy. We start by showing the results from contraction and re-ranking experiments, in the spirit of Kleinberg et al. (2018). Through these exercises we estimate the increased effectiveness that could be attained by excluding some current beneficiaries that are not ML targets, and (under the assumption of selection on observables) by substituting them with firms that are not treated under the GF rules, but that should have been eligible for the collateral according to ML. Next, to relax the selection-on-observables assumption, we exploit the threshold for assignment implied under the GF rules and run a Regression Discontinuity Design (RDD) experiment (as in de Blasio et al., 2018), separately by ML-targeted and non ML-targeted groups of firms. We find that effectiveness is higher for the firms identified by ML as targets.

We show that around 47 per cent of the resources currently allocated by the GF rule go to firms that are not a target according to our ML algorithms. By channeling these resources to other firms identified as ML-target, the effectiveness of the policy improves significantly.

One pitfall of our approach is that we train the ML algorithm by using data for a period in which the guarantee was already available. While this is a rather common situation for policymakers who try to re-design a scheme that is already in place, our ML prediction is likely to show a higher out-of-sample (forecasting) error with respect to the case in which a policy is yet to be introduced. Our results should therefore be taken as a conservative estimate of the benefits that could be obtained by using ML instead of the naïve rules. If the data were not contaminated by previous treatment, the prediction would have been more accurate and the gains from ML even larger. We discuss the importance, in our case, of other issues that are typically related to the use of ML for policy decisions, such as transparency and omitted payoffs. We show that our preferred random forest algorithm is the one that performs worse on transparency grounds. However, it is not clear the extent to which off-the-shelf alternatives, such as decision-tree and LASSO routines, improve the transparency of the assignment process, and whether the GF rule, based on a scoring system that uses as input balance-sheet data, can be considered superior when accountability is at stake. We also argue that potentially omitted payoffs for the policymaker might derive from the allocation of the guarantees across banks and territories. To

this aim we contrast the distribution of collaterals across lenders and areas that would derive from ML targeting with the one based on the GF naïve rule.

While the literature on ML for policy analysis is now booming (see Athey, 2018, for an updated review), the papers that deal with ML techniques to tailor the assignment of a policy are few. McBride and Nichols (2015) propose to use ML to improve poverty targeting. Andini et al. (2018) exploit ML to show how to re-target a scheme intended to boost consumption. As for the literature on credit guarantees, the paper closest to ours is Riding et al. (2007), which focuses on measuring additionality in a standard (non-ML) econometric framework.

Apart from the tailoring of the policy, we also make other contributions to the literature on credit constraints and default risk assessment. As in Jiménez et al. (2012, 2014), we use a measure for credit constraints that exploits some unique features of the CR, which is collected by the Bank of Italy acting in its capacity as bank supervisor. The register records monthly information requests lodged by banks on borrowers (which are currently not borrowing from them). As the CR database also contains detailed monthly information on all, new and outstanding, loans, we can match the set of corresponding loan applications with the actual variation in bank credit granted to the applicant over the following months. We therefore provide a prediction of credit constraints which is based on hard data from a sizable dataset. To the best of our knowledge, no previous forecasting exercise has ever been attempted for an indicator of credit market access. The second leg of our ML exercise predicts non-performing loans. In this case, the forecasting industry is developing ML techniques and we have the advantage of running this exercise on a very large and reliable database. All in all, we believe that our predictions might be useful for a broader audience that includes bank supervision agencies, place-based policy agencies (for instance, the EU authorities) who try to channel resources towards areas that suffer from lack of access to the credit market, and even commercial banks, who might be interested in predicting the likelihood of repayment across their customers.

The remainder of the paper is structured as follows. Section 2 discusses the rationale of public guarantee schemes and highlights some concrete examples of programs around the world. Section 3 describes the characteristics of the Italian scheme, the Guarantee Fund. In Section 4 we discuss how we train the ML algorithm to predict firms that are both credit constrained and creditworthy, and we compare the algorithm predictions with the naïve GF rule. In Section 5 we provide evidence that ML targeting ensures higher program effectiveness. Section 6 discusses some pitfalls of our strategy, including the issues related to omitted payoffs and the transparency problem. Section 7 concludes. We also include a detailed Appendix, which discusses the more technical aspects of the ML algorithms.

2. Credit guarantees: additionality and financial sustainability

The impact of credit guarantee schemes depends on whether they actually reach firms that are credit constrained. However, reaching credit-constrained firms involves greater risk taking and hence a greater probability of incurring financial losses, putting at risk the guarantee schemes' financial sustainability. As stated by the World Bank (2015), "it is essential that credit

guarantee schemes be properly designed and operated to achieve both outreach and additionality in a way that is financially sustainable". In the concrete experience of policy making, this has been a challenging task. While the financial sustainability of the schemes can be approximated by means of firm risk screening models (i.e. credit scoring models), a guidance to reach additionality based on a measure of the firms' credit constraints is mostly lacking. As a result, as argued by Zia (2008), credit guarantees usually fail to reach constrained firms.

There are, however, some notable exceptions where the presence of credit constraints is explicitly addressed within the credit guarantee scheme. The U.S. Small Business Administration (SBA), for instance, requires firms applying for a guaranteed loan to demonstrate the inability to obtain credit available elsewhere on reasonable commercial terms without the SBA guarantee (so-called "credit not available elsewhere" test). This task primarily involves the potential lender, which must specify which factors prevented the financing from being accomplished without SBA support and include the explanation in the applicant's file.[3] Other exceptions are FAMPE (Fundo de Avail às Micro e Pequenas Empresas) in Brazil, and KGF (Kredi Garanti Fonu) in Turkey. In both cases, the guarantee fund supports SMEs that are creditworthy but proven to lack sufficient collateral. A second, related, approach to reach credit-constrained firms involves limiting the guarantee programs to specific categories of firms or sectors, for which there is clear evidence of problems in accessing credit (i.e. firms operating in poorer areas, start-ups, female entrepreneurships). Hence, with respect to the previous approach, the inability to obtain market credit is assessed at an aggregate rather than individual level.

While in principle both approaches provide a possible solution to increase additionality, they are not free of drawbacks. With respect to the first approach, some requirements, as claimed by Vogel and Adams (1997), might be circumvented by manipulation and increased "verification" cost (Beck et al., 2008). As for the second, when the financial difficulties are assessed at the aggregate level, there will still be credit-constrained and creditworthy firms that are left without funds (i.e. firms belonging to non-targeted sectors or areas; see Deelen and Molenaar, 2004) and firms from the chosen sectors and areas that are financially unconstrained and thus receive an undeserved benefit.[4]

[3] Acceptable factors include, among others: the requested loan has a longer maturity than the lender's policy permits; the requested loan exceeds either the lender's legal lending limit or policy limit regarding the amount that it can lend to one customer; the collateral does not meet the lender's policy requirements; the lender's policy normally does not allow loans to new businesses or businesses in the applicant's industry (see https://hcdc.com/credit-elsewhere-test/). In the new Standard Operating Procedure 50 10 5(J) which took effect on 1 January 2018, the agency has issued some further guidance on how to identify that credit is not available from private sources.

[4] Our ML targeting focuses on the assignment rule. Note, however, that also the pricing and the coverage ratios of the guarantee have been used to enhance the capability of reaching financially constrained firms. For instance, Honohan (2010) argues that low fees may lead to the provision of guarantees to firms that are not financially constrained. When fees are high enough, only credit-constrained firms should be encouraged to apply to the scheme. However, as argued by Saadani et al. (2011) and the OECD (2013), high fees could also lead to adverse selection, with the more risky borrowers taking part in the scheme. Setting a higher coverage ratio for more risky firms is another approach that has been widely adopted in order to increase additionality. This approach relies on the fact that banks generally require higher coverage to provide loans to riskier firms, typically seen as more credit constrained (Saadani et al., 2011). Although this approach is widely used, it also displays some weaknesses. While a high coverage ratio *per se* does not prevent unconstrained firms from applying for the guarantee, it could lead to moral hazard behavior from both firms and banks (see, among others, Uesugi et al., 2010).

On the whole, a reliable mechanism for predicting which firms should be considered under a scheme of public collateral is still lacking. Our benchmark assignment mechanism, based on ML techniques, might be seen as a further step in tackling the difficulties in accessing credit around the world.

3. The Italian Guarantee Fund

The Italian GF started its activity in 2000. Initially the volume of bank loans with public guarantees was quite small, totaling €11 billion until 2008. With the advent of the crises it experienced a boom. From 2009 to 2016, €86 billion in loans to SMEs benefited from the public guarantee. The growth in volumes reflects the desire of the Italian authorities to counterbalance the effect of the credit crunch. The latter was particularly severe for SMEs that, in an environment of increased credit risk, experienced a more significant drop in credit flows and a stronger rise in interest rates with respect to larger firms (Ministero dello Sviluppo economico, 2015; Comitato di gestione del Fondo di garanzia, various years).

The provision of GF guarantees is limited to SMEs, defined according to EU criteria, in the private sector, which includes manufacturing, construction and services. However, some specific sectors, such as agriculture, automobile and financial services, are not covered by the scheme because of the limitations imposed by the EU regulation on competition. The public guarantee insures up to 80 per cent of the value of a bank loan. For each firm, however, there is a maximum amount of guarantee, which is equal to €1.5 million. The GF can guarantee both short-term and long-term loans and there are no constraints in terms of the final use of the funding by the borrower. It is important to notice that, in case of default, the financing institution can immediately call on the GF to meet its obligation ("first demand guarantee").

According to the GF procedure, a SME that needs to borrow might ask the bank to apply for a public guarantee (alternatively, it is the bank that might propose to the firm to apply for the guarantee). The bank has to verify the eligibility of the firm for the scheme through a scoring system (a software) provided by the GF. The scoring system is designed to minimize the likelihood that a firm defaults on its debt; no consideration is given to the actual financial constraints of the SME. The scoring system takes into account four indicators (that slightly differ according to economic sector) of the firms' financial condition in the two years preceding that of the application: the soundness of firms' financial structure is measured for the industry (service) sector by the ratios of equity and long-term loans to fixed assets (short-term assets on short-term liabilities) and equity to total liabilities (short-term assets to sales); short-term financial burden is measured by the ratio of financial expenses to sales; cash flow is measured by the ratio of cash flow to total assets. For each of the two years preceding that of the application, the software calculates from the values of the balance-sheet variables a single "partial score". The partial score is collapsed into three categories (A=good, B=intermediate, C=bad), as described in de Blasio et al. (2018). The combination of the two partial scores, one for each year (with higher weights envisaged for more recent scores), allows the assignment of the final score. According to the final score, the applicant firms are split into three types (0, 1, and 2). Type-0 firms are not

eligible. Type-1 and Type-2 firms are both eligible but do not automatically receive the treatment. They have to go through a further assessment, which is more demanding for the Type-1 firm, as they have worse scores (i.e., poorer lagged balance-sheet observables).[5] The additional assessment concludes with final approval or rejection. Rejection, however, has been a rare event (3.8 per cent of the applicant firms were rejected over the 2011-16 period).

In December 2017 the Italian Guarantee Fund was the subject of a reform, primarily aimed at: (i) enlarging the number or potential beneficiary firms, (ii) improving the screening of firms to exclude those that are not creditworthy, and (iii) increasing the support to creditworthy firms that are more exposed to the risk of being credit constrained. The central point of the reform was the adoption of a new rating model to assess the creditworthiness of the firms, based on a larger set of information with respect to the mechanism described above. In order to tackle credit rationing, the new rules allow more risky (but still creditworthy) firms to benefit from a larger share of the loan covered by the guarantee. The reform is still not operative at the time of writing. It is expected to become operative by end 2018. The analysis carried out in this paper is hence solely based on the pre-reform GF rules.

4. The prediction problem

We now illustrate how to design an allocation rule to identify, on the basis of observable pre-determined characteristics, which firms are more likely to be both financially constrained and creditworthy. In order to identify these firms, we estimate two separate ML predictive models for the two conditions. Being financially constrained is measured through the discrepancies between firms' credit demand and its supply, which are neatly traceable using the Bank of Italy's Credit Register (CR). Being creditworthy is proxied by the occurrence of adjusted bad loans (see Subsection 4.2 for the definition), which is also observable in CR. The next Subsection illustrates the ML algorithms. Subsection 4.2 presents the data. Subsection 4.3 discusses the prediction results (more details about how we practically implement and estimate each algorithm are reported in the Appendix). Subsection 4.4 contrasts ML targeting with the GF assignment rule.

4.1 ML algorithms

In our exercise we estimate two separate predictive models for being credit constrained and creditworthy, that is:

$$credit\_constrained_i = f(X_i) + \epsilon_i \qquad (1)$$

$$creditworthy_i = g(X_i) + \eta_i \qquad (2)$$

where i indexes the loan application from a firm in a given quarter, X_i is a set of P observable characteristics for the firm at the time of the application, $f(.)$ and $g(.)$ the two functions to be

[5] According to the GF guidelines, the additional assessment is referred only to cash-flow requirements for Type-2 firms. As for Type-1 firms, the additional assessment is an in-depth analysis of the economic and financial situation of the firm. Again, the aspects related to credit constraints do not matter.

learnt from the data, ϵ_i and η_i are noise. The outcomes are two binary variables $credit\_constrained_i$ and $creditworthy_i$ assuming value one if the application belongs to the respective status. Estimating separate models does not imply that we are assuming that the two events are statistically independent. In Subsection 4.3 we show the relation between the two predictions and we discuss the implications for our analysis. From an econometric perspective, our purpose is to predict each status using the same set of observable characteristics. One could think of our prediction problem as a system of simultaneous equations where each status depends on both the covariates and the (true) probability of the other status. This would boil down to two equations where the right-hand sides are a function of the observable characteristics. However, we do not observe the latent probability of each status and therefore such an approach is unfeasible. Another unexplored issue is that, by modeling the correlation between the error terms of the two equations, one could improve the prediction for the joint status (credit constrained and creditworthy). Finding a solution to this issue in the ML context is not straightforward. An alternative could be to predict directly the target firms as those that are both constrained and creditworthy. In this case, however, our data are such that the target vs non-target status is almost only informed by the credit-constrained status rather than the creditworthy one, as the latter is highly unbalanced towards the creditworthy status. In addition, no improvement has been reached in terms of misclassification error. Further details are provided in the Appendix.

As we do not know the true functions $f(X_i)$ and $g(X_i)$, our aim is to estimate (or train, in ML jargon) them by using a model that has good forecasting performance out-of-sample, because the rule is meant to be used for future assessments of new requests for the GF guarantee. In this respect, ML tools are particularly useful (Mullainathan and Spiess, 2017) as they aim to minimize out-of-sample forecasting error. In short, such tools rely on highly flexible functional forms, where greater complexity in the model improves its in-sample fit but reduces the out-of-sample fit of the selected model. The complexity of the model is set through a regularization parameter, chosen by cross validation in order to minimize the out-of-sample error (Hastie et al., 2009). As a criterion for selecting the best complexity parameter and the best model across different alternatives, we look at the misclassification rate, which is the fraction of observations that are predicted to belong to the wrong class. Unlike in standard econometrics, ML models do not focus on obtaining unbiased estimates of the two functions, but rather on minimizing the out-of-sample forecasting error. Their objective function therefore allows for some bias in the estimator if this reduces the variance of the prediction.

In practice, we employ and compare three different ML algorithms, the decision tree, the random forest and the logistic LASSO regression. Before fitting our models, we randomly split our sample into two subsamples, a training sample and a testing one, following the 2/3 – 1/3 division rule (as suggested in Zhao and Cen, 2014). We then fit our models on the training set and test their out-of-sample predictive performance over the testing set. In what follows we introduce the three algorithms for readers who are less familiar with ML. In Appendix A.3 we discuss the details of their implementation, including our strategy for dealing with the

unbalancedness in the creditworthy status (as most of the observations are such that $creditworthy_i = 1$).

The decision tree is a classification algorithm that provides the researcher with a clear scheme (the tree) to follow for targeting. Intuitively, the decision tree divides the set of possible values of all the variables into J non-overlapping regions $j = 1, \dots, J$. At step 1, starting from the whole sample, the algorithm identifies the variable x_{pi} from X_i and the threshold s_1 such that, by splitting the sample into two regions $x_{pi} < s_1$ and $x_{pi} \geq s_1$, we obtain the highest reduction in the sum of the Gini impurity index across the two regions.[6] At each subsequent step, the tree continues splitting the sample by finding a variable and a threshold that lead to the highest reduction in the impurity index. The tree can be grown as long as there are at least some observations in each node. However, a high number of levels in a tree (i.e., a very complex tree) is likely to overfit the data, leading to poor out-of-sample predictions. By setting a regularization parameter c_p, it is possible to reduce the complexity of the tree (see Hastie et al., 2009). Formally, the tree choice solves an optimization problem:

$$min_T \sum_{l=1}^{|T|} N_l L_l(T, y_l) + c_p|T| \qquad (3)$$

where *T* is the tree used to forecast the status *y*, $|T|$ is the total number of leaves, l is a leaf of tree T, N_l is the number of observations in the leaf, $L_l(T, y_l)$ is a loss function (the Gini impurity index in our case), and y_l is the vector of outcomes for observations in the leaf. Setting a low c_p would lead to a large tree with a good fit in the training sample, but possibly with large out-of-sample error. By setting a higher c_p we reduce its size (we "prune" the tree) and therefore we reduce the risk of overfitting. We set the complexity parameter by 10-fold cross-validation, trying to minimize the out-of-sample misclassification error. Instead of choosing the parameter that reaches the minimum cross-validation error, we use a rule-of-thumb, common in the ML literature, which takes the smallest c_p whose associated error is larger than the minimum cross-validation error plus its standard deviation.

The random forest algorithm provides an improved prediction by averaging the classification produced by *n* decision trees. Each tree is estimated on a new sample bootstrapped from the original training, but allowing only for a (randomly drawn) subset *m* of the *P* predictors. Each tree is grown to its maximum extension, without pruning (and therefore without setting an optimal c_p). This procedure aims to reduce the correlation between the trees, in order to reduce the variance of the prediction. Again, in order to optimally define the parameters on which the algorithm is based, a 10-fold cross validation is performed to select the two parameters *n* and *m*.

The logistic LASSO algorithm provides a prediction that is based on a logit model (with a linear index) where the estimated coefficients are penalized according to their magnitude. To allow for non-linearities in the logit index, the vector of covariates is expanded to $\widetilde{X}_i$ of dimension $\tilde{P}$ by including pairwise interactions between the variables in X_i for observation i in

[6] For each region, the Gini impurity index is equal to $2f(1-f)$ where f is the fraction with the outcome equal to 1 (that is the fraction belonging to the status).

the training sample ($i = 1, \dots, N$). LASSO solves the following optimization problem (Hastie et al., 2009):

$$\max_{\beta_0,\beta}\{\sum_{i=1}^{N}[y_i(\beta_0 + \beta'\widetilde{X}_i) - \log(1 + e^{\beta_0+\beta'\widetilde{X}_i})\,] - \lambda\sum_{j=1}^{\tilde{p}}|\beta_j|\} \quad (4)$$

where λ is a penalization parameter, β_0 is a constant and β' is a transpose vector of the β coefficients to be estimated (together with the constant). The penalization implies that only a subset of indicators will have coefficients other than zero. It is crucial to choose λ, again through 10-fold cross validation.

4.2 Data and sample selection

While we cannot directly observe when a firm makes a bank loan application, we can draw information from the Bank of Italy's CR.[7] In particular, we use as our main data source the requests of preliminary information (PI) collected by the CR. The PI request is an instrument used by banks to gain information on the reliability of new potential borrowers: through a PI request, banks can obtain detailed information on the credit history of their loan applicants.[8] Given that obtaining information through a PI request is not free of cost, it is reasonable to assume that the decision to bear the cost of inspecting firms' credit history always follows the presence of a previous loan application by the inspected firms to the PI-requiring bank. Throughout the paper we will therefore treat each PI request as a loan application and we will use the two terms interchangeably.

Using the PI requests we build two datasets consisting of Italian limited companies that applied for a bank loan (i.e. firms for which banks issued a PI request) in 2011 or in 2012. We chose 2011 and 2012 as sampling years because they leave us with a good number of follow-up years, while still allowing us to draw information on firms' past history over two years. From the dataset we exclude (i) firms for which we do not have balance-sheet information on the two years preceding the PI request; (ii) firms that have never had lending relationships with banking institutions in the two years preceding the PI request. These firms, which include for instance start-ups, are likely to be fundamentally different, and therefore we would need to devise a separate forecasting and assignment exercise. This lies outside the scope of our paper. Our algorithm and prediction exercise is therefore limited to those firms that have standard balance sheets and for which there is sufficient history in CR.

The final sample on which we train and test our ML algorithm is composed of nearly 190,000 firms that made a bank loan request in 2011. This sample is randomly split into a 2/3 training sample (for estimating the models) and 1/3 testing one (for validating and comparing the models). Firms that applied for a bank loan in 2012 constitute instead our hold-out sample (see the Appendix for more details on this sample), which we will use in Subsections 4.4 and 5 to

[7] A similar register is maintained by the Bank of Spain (Jiménez et al., 2012 and 2014).

[8] The CR retains information at the loan level on all loan contracts granted to each borrower whose total debt from a bank is above €30,000.

compare the current GF assignment rule with the one that we devise based on the ML algorithms.

Our sample is likely to represent only a subset of the total number of firms that request credit. In fact, it excludes those credit-requiring firms for which PI request is not issued. This is the case when the firm that applies for a loan is already known to the bank, or the firm is outstanding and no further screen is needed by the bank. The presence of credit relations for which a PI has not been issued will likely drag our sample towards less financially sound firms. However, this is not an issue for this study, as firms that are indeed financially stable are not the primary target of the GF in first place.

Firms may issue more than one loan request over time, even within the same year. Given that our algorithm should be able to forecast the firm's condition at the time of the application, we also exploit the information from multiple PI requests during 2011. In our sample, therefore, each firm may appear more than once and each observation is a single PI request. However, given that most of our explanatory variables change only at the quarterly level, we assume that different loan applications issued by the same firm within the same quarter refer to the same project. Different loan applications by the same firm in different quarters are instead considered as separate observations and included in the sample. The final 2011 dataset is therefore composed of 278,355 observations. Approximately 2/3 of this dataset pertains to the (randomly selected) firms of the training set (185,256 observations, relative to 123,276 firms observed on average for 1.5 loan applications), while the remaining 1/3 pertains to the test set (93,099 observations, relative to 62,052 firms observed on average for 1.5 loan applications).

For each firm that applied for a loan in a given quarter of 2011 we devise two outcome variables, which will be the object of our learning classification algorithms. The first is an indicator of whether the firm is credit constrained, namely if its total amount of granted loans has not increased six months after the PI request.[9] In our sample about two thirds (66.2 per cent) of loan applications refer to firms that are credit constrained, a figure in line with those obtained from the Survey on SMEs access to finance (ECB, 2015). The second is an indicator of whether the firm is creditworthy, namely if it does not have "adjusted bad loans" in the three-year window following the PI request.[10] About 86 per cent of the applications refer to firms that are creditworthy.

The forecasting exercise is run using a set X_i of explanatory variables that are observable by the policymaker (in our case, the GF) at the time she is required to assign the guarantee according to the decision rule in place. We focus on CR data on lending from the banking system

[9] The measure considers the total amount of bank loans, and not just the loans granted by the banks that issued the PI request about the firm, in order to control for those cases where the credit is issued to the firm by banks not requiring a PI. For more details on the credit-constraints index based on PI requests, see Albertazzi et al. (2017), Carmignani et al. (2017), and Galardo et al. (2017).

[10] A firm has adjusted bad loans if it is reported as insolvent by a bank that accounts for at least 70 per cent of the firm's total bank loans, or if it is reported as insolvent by two or more banks that together account for at least 10 per cent of the firm's total bank loans. See "sofferenze rettificate" at: https://www.bancaditalia.it/footer/glossario/index.html?letter=s.

and balance-sheet information from the Cerved database. In both cases, we include not only variables in levels, but also a measure of their change over time. We also include some additional variables capturing firm-specific characteristics: firm age, location and sector indicators. Finally, we introduce a dummy variable that takes the value of one for firms that have already been beneficiaries of the GF program in the years preceding the PI request (data on GF beneficiaries have been available since 2005). The complete set of covariates includes 108 variables[11] (see Table A1 in the Appendix for the complete list and a brief description; Table A2 provides summary statistics). In order to minimize information redundancy (Fan and Lv, 2008), we submit our covariates set to a pre-processing procedure before applying ML techniques in the two predictive exercises (see the Appendix). We stress the fact that these variables are currently accessible to the GF administrators who are already employing them and, more in general, to the policymaker. In terms of information requirements, therefore, we are not imposing any additional burden.

4.3 Prediction results

We carefully assess the prediction performances of the three ML models (decision tree, random forest and logistic LASSO) in the Appendix. Combining the evidence, our preferred model is random forest. Figure 1, panels a and b, shows that the predicted probability of belonging to each status (i.e. being creditworthy and credit constrained) is strongly correlated with the actual rate. Figure 1c shows that being creditworthy is correlated with being credit constrained, but there is large dispersion around this relation. The predicted probability of being constrained initially strongly declines with the probability of being creditworthy, as expected given that less risky firms are more likely to get access to credit. But, in the rest of the distribution, the relation between the two probabilities is flatter. Furthermore, there is a large dispersion around each point. Hence a measure of default probability does not seem to provide enough information to also target firms that are credit constrained.

One concern is how it is possible that two firms with the same risk are not equally credit constrained. There are three possible explanations for this. The first is that there is true heterogeneity in credit constraints even for firms with the same risk, thus it makes sense to further exploit the information on predicted credit constraints to better target the policy. This true heterogeneity in constraints can be due to the presence of other guarantees or different forms of collateral (which affect the banks' loss given default) that we are not able to directly observe and that could not be included as part of the assignment rule (as an applicant might simply not declare them). A different level of credit constraints for firms with the same risk might also depend on banks' policies on risk diversification and on the amount of delegation in credit management, which could impose limits on specific firms, sectors or territories. The second explanation is that banks have more information than we do and, therefore, assess risk better. For any given probability of default predicted by us, some firms are actually more risky (and the

[11] We recover the same set of observables for firms that applied for a loan in 2012, which form our "hold-out" sample.

banks know that), hence selecting as eligible only the credit-constrained firms may lead the GF to get the lemons.[12] This issue might hinder the ability of ML-targeting to improve effectiveness, and calls for the ex-post evaluation that we provide in Section 5. A third explanation is that our finding of dispersion in credit constraints is only due to measurement error. In this case, though, we should find no improvement in terms of additionality, hence, once again, the issue boils down to the ex-post evaluation in Section 5.

Finally, looking at the two dimensions is extremely important even if one believes that the credit-constrained status is a deterministic (or quasi-deterministic) function of risk. The policymaker may want to design an eligibility rule that prioritizes firms at the level of risks that are associated with more constraints, for instance, because she believes these constraints exceed the social optimum. To avoid doing so arbitrarily, we still need to empirically evaluate the relation between the creditworthy and the credit-constrained status, hence our results might prove valuable for this purpose.

We therefore combine the two models to look at our final target, i.e. those firms that are predicted to be both creditworthy and credit constrained. These are the firms whose forecasted probabilities for the two conditions are both larger than 0.5. For this joint status, the misclassification error, reported in Table 1, is 36.8 per cent.

If we were to use this forecasting as a condition of eligibility for the GF, we could therefore wrongly allow access to the guarantee to three groups of misclassified firms: (i) constrained but not creditworthy; (ii) not constrained but creditworthy; (iii) neither constrained nor creditworthy. In terms of the financial stability of the Guarantee Fund, the most problematic here are groups (i) and (iii), as they have a low likelihood of paying back the loan. Fortunately, misclassification is largely concentrated in the less problematic group. Indeed, among the 22,779 misclassified observations, 73.7 per cent belong to group (ii), 20 per cent belong to group (i) and 6.3 per cent belong to group (iii).

4.4 ML rule vs GF rule

Let us assume that we have to decide whether or not to provide the public guarantee based on the observables available in 2012. We want to compare the GF rule, which evaluates whether a firm is eligible or not on the basis of the economic and financial indicators described in Section 3, with the ML rule. We consider only firms belonging to the sectors that are currently eligible for the GF scheme.[13] In order to evaluate whether a firm is eligible for the Fund guarantee, we apply the GF scoring procedure to the firms in our dataset. It is worth noticing that we would fail to exactly replicate the GF outcome in 100 per cent of the cases, as: (i) we do not

[12] However, the opposite might also be true if weaker banks misallocate credit towards firms on the verge of bankruptcy (Schivardi et al., 2017) or if banks favor connected firms (Barone et al., 2016).

[13] The sectors eligible under the GF scheme are divided into two groups: manufacturing, construction and fishing (which we will refer to as group1) and the tradable sector, hospitality industry, transportation and other private service sectors (group 2). Firms belonging to these two groups face a slightly different screening procedure by the Guarantee Fund, based on a different set of balance-sheet indicators (and scoring thresholds).

have access to the firms' original balance sheet data that were provided to the Fund but, instead, we observe less detailed reclassified balance sheet data (drawn from the Cerved archive); (ii) we do not observe when the request for the guarantee was issued. Hence, we do not precisely know which are the two years that should be considered to compute the GF scores. In order to obtain some guidance, we consider 2011 and 2010 balance sheet data for those firms for which banks have issued a PI request in the last quarter of 2012; we consider 2010 and 2009 balance sheet data for the other firms. Notwithstanding these difficulties, we replicate the Fund eligibility mechanism fairly well (Table 2). Only about 2.3 per cent of the firms that received the Fund guarantee in 2012-13 are classified by us as not eligible when we replicate the GF rule on reclassified Cerved balance sheet data.

Table 3 compares the GF rule with the ML one. Overall, the ML rule is more selective with respect to the actual Fund rule. Out of roughly 90,000 firms in our dataset, about 80 per cent of them would be selected by the ML targeting mechanism, while about 95 per cent are eligible according to the GF rule. In particular, the ML targeting would exclude about 20 per cent of the firms that are eligible according to the GF assignment mechanism (considered "outstanding" or "fair" firms: see Section 3). On the other hand, the ML rule would select about 75 per cent of the firms that are not eligible ("poor" firms) according to the GF rule. This evidence is in line with the rationale of the ML algorithm, which grounds eligibility on both creditworthiness and the actual need for external funds. As a result, GF eligible firms which have fair access to credit are not targeted by ML; on the other hand, firms that have low capacity to access credit, while still being creditworthy, are targeted by ML. Table 4 shows in detail the characteristics of the 16,860 firms that are eligible according to the GF but not selected by the ML algorithm. About 70 per cent of these firms are creditworthy but not constrained; about 25 per cent are constrained but not creditworthy, while only 5 per cent are neither creditworthy nor constrained.

In order to shed more light on the differences between the GF eligibility mechanism and our ML targeting rule, we consider the full set of about 90,000 firms in our dataset and estimate a simple linear model where the dependent variable y is a dummy taking value 1 if the firm is eligible according to the GF scoring mechanism and 0 otherwise. Our independent variables are: the ML predicted probability of being creditworthy; the ML predicted probability of being credit constrained; a dummy equal to 1 if the firm belongs to the sectors of manufacturing, construction and fishing and 0 if the firm belongs to the tradable sector, hospitality industry, transportation and other private service sectors. The results in Table 5 show a positive and statistically significant correlation between the probability of being eligible for the GF and that of being creditworthy. On the other hand, being credit constrained is negatively correlated with the probability of being eligible. These results strengthen our claim that the GF eligibility rule is largely based on firms' creditworthiness, while it overlooks whether a firm is credit constrained or not.

In Figure 2 we compare our ML predictions with a continuous version of the GF eligibility score *s* (see Subsection 5.2), which is a normalized measure of the distance to each value of it. The basic criteria for GF eligibility are met when *s* crosses the 0 threshold. Given that

the GF scoring procedure essentially refers to the financial soundness of the firm, as expected the eligibility score is positively correlated with our ML predicted probability of being creditworthy and negatively with the one of being credit constrained (panels a and b). Yet, even for high values of the GF score there is a sizeable share of firms that are not predicted to be creditworthy according to ML, and vice versa. If we look at the predicted joint status of being worthy and constrained (panel c), we note that the association with the eligibility score is quite flat. Hence there are several firms that would meet both requirements of ML targeting but which are far from the GF eligibility conditions.

5. Evidence on ML-targeting effectiveness

We now assess whether replacing the actual GF eligibility rule with our benchmark ML-based assignment mechanism increases the impact of the policy. We start (Subsection 5.1) by showing the results from contraction and re-ranking experiments, in the spirit of Kleinberg et al. (2018). Next, in Subsection 5.2, we exploit the threshold for assignment implied under the GF rules and run a RDD experiment (as in de Blasio et al., 2018), separately with reference to the ML-targeted and non ML-targeted groups of firms.

5.1 Contraction and re-ranking

In order to assess whether the ML-based targeting rule leads to an improvement in the effectiveness of the Fund, in this Subsection we rely on two different strategies: crude comparisons and matching. In the first one, we focus on the subset of firms that received the GF guarantee in the period 2012-13. We split these firms into two groups, according to whether they are targeted by ML or not. Then, we compare the average observed performance of the two groups with respect to a set of variables measuring both financial and real outcomes over the period 2011-15. The idea behind this exercise is that, if the group of ML targeted firms performs better than the other group, the policymaker might increase the effectiveness of the policy by simply excluding a subset of the Fund's eligible firms. This approach is referred to as a contraction experiment in Kleinberg et al. (2018). It is very straightforward, as it relies on a simple comparison on observed average outcomes. In our sample of roughly 90,000 firms, about 7,000 firms received the Fund guarantee in the years 2012-13. Among them, about 4,000 firms (60 per cent) are also selected as target by the ML algorithm, while 2,869 beneficiary firms are not selected (Table 6). Among the latter, about 70 per cent are discarded because they are not ML predicted as credit-constrained firms.

We observe the performance in the post-treatment years of the two groups of treated firms: those selected by the ML algorithm vs those that are not. The subset of treated firms selected by ML (4,042 firms) performs much better than those that would have not been selected by the ML (2,869) across all the variables considered (granted loans, number of banks, adjusted bad loans, fixed assets, sales; Table 7). For instance, the ML targeted firms displayed a cumulative growth rate of granted bank loans of about 8 per cent, while the non-targeted group

displayed negative growth of about 48 per cent. The targeted group also performs much better when it comes to adjusted bad loans, with 2015 rates being half those of the non-targeted group. Apart from the growth of fixed assets and profitability in 2015, the differences in the performances of the two groups of firms are always statistically significant. This evidence suggests that using our ML algorithm simply as a further screening device for all the firms previously selected by the Fund's rule ("contraction") would considerably improve the overall performance of the policy.

The use of ML in a "contraction" approach would systematically lead to a drop in the number of beneficiary firms, probably turning ML into a not-so-appealing tool for a policymaker. In particular, no evidence is given about the effects of replacing some of the beneficiary firms – those not targeted by ML – with some non-beneficiary ones which, however, are targeted by ML. To shed some light on this we undertake a second experiment (this is referred to as "re-ranking" in Kleinberg et al., 2018), which relies on the selection-on-observables assumption. We use standard matching procedures to associate any of the firms that are targeted by ML but did not get access to the Fund to a companion firm chosen among those that receive the collateral and are ML targeted. We then impute the performance of the latter to the untreated ML-targeted firms. Finally, we compare the imputed performances of untreated ML-targeted firms with the observed performances of treated firms that are not ML targeted.

Operationally, we start by considering the set of about 3,500 firms that did not receive the collateral because they were not eligible, but are ML targeted (see Table 3). The matches are selected among the set of about 4,000 ML-targeted and beneficiary firms. Matching is implemented through a nearest-neighbor algorithm, which uses the following variables: ML predicted probability of being creditworthy, ML predicted probability of being credit constrained, as well as a set of covariates measured in in the pre-treatment year such as granted bank loans, disbursed bank loans, assets, sales, number of banks lending money to the firm and firm's economic sector. We only keep firms that have a reasonably good match, leaving us with a final sample of 1,303 firms.[14] Table 8 describes the average performance of this group of firms, across a set of outcome variables. If selected and admitted to the treatment, these firms would have done significantly better than the group of 2,869 treated but non ML-targeted firms. The policymaker could improve the overall performance of the Fund by reallocating the collateral according to the ML suggestions.

We now turn to the gains that the alternative targeting rule would yield in terms of public resources allocated to the policy (Table 9). Of the 6,911 beneficiary firms, 41.5 per cent of them are not targeted by ML. When we consider the amount of public funds, the extent of misallocation is even larger: the resources granted in the form of public guarantees to firms that are not targeted by ML reach 46.5 per cent. On average, treated firms not targeted by ML are

[14] We perform a nearest-neighbor matching using the Stata routine *nnmatch* (Abadie et al., 2004). We include the matches whose distance falls within the standard stopping rule (4 times the mean increase in the average distance of the first quartile of distance).

characterized by larger public-guarantee backed loans and larger guarantees (+26 per cent and +22 per cent, respectively).

Treated firms not targeted by ML can be of three types: (1) not constrained and creditworthy (60 per cent of 2,869 firms); (2) constrained and not creditworthy (30 per cent); (3) not constrained and not creditworthy (10 per cent). The guarantees are mostly channeled to the first type of firms, attracting about 70 per cent of the financed amounts. Hence, the bulk of guarantees are directed towards firms that have, presumably, a good capacity to access credit. While, in this case, the risk of not recovering the public guarantee is rather low, the deadweight loss refers to the circumstance that the public collateral could have been used for firms that face credit access difficulties. The remaining 30 per cent concerns firms that are not creditworthy (whether constrained or not constrained). For these firms the consequences of mis-targeting are more serious and potentially very onerous for the public finances. All in all, if the guarantees channeled to the 2,869 treated firms not targeted by ML were, instead, used to guarantee new (ML-targeted) firms (keeping the same average guaranteed amount characterizing the group of ML-targeted beneficiary firms) the number of beneficiary firms would increase by about 630 (about 3,500 new firms would receive the GF guarantee; 2,869 currently beneficiary firms would instead be excluded), leading to an increase in the total number of beneficiary firms of about 10 per cent.

5.2 Evidence from RDD

In this Subsection we further investigate whether using the ML-based assignment mechanism, rather than the GF current rule, leads to an improvement in the effectiveness of the GF by exploiting a more credible identification strategy with respect to matching and relaxing the assumption of selection-on-observables.

We follow de Blasio et al. (2018) and exploit the GF eligibility mechanism to estimate the impact of the guarantee using a fuzzy-RDD strategy, which allows for imperfect take-up of treatment. As in their case, the compliance is imperfect both below and above the eligibility threshold. Above the threshold, we have eligible firms that have not applied to the GF, and eligible applicant firms eventually rejected by the GF. Below the cutoff, noncompliance is associated with the fact that we fail to successfully predict the eligibility status for firms using balance sheet data. The fuzzy-RDD identification critically rests on a discontinuity of the probability of treatment at the threshold, as well as on the absence of manipulation of the assignment variable (see Lee and Lemieux, 2010).

We follow de Blasio et al. (2018), relying on a continuous forcing variable that builds on the measure of eligibility applied by the GF. With respect to them, we innovate along two dimensions. First, as we consider a different treatment-period window (2012-15, basically subsequent to theirs), we offer a new test of the effectiveness of the GF. Second, and most relevant to our study, we assess the impact of the effectiveness of GF separately for firms that are targeted or not by ML. A greater impact of the GF for the subgroup of ML-targeted firms would strengthen the previous results.

This analysis is conducted over a sample of about 63,000 firms, which might have benefited from the GF between 2012 and 2013 and for which we are able to observe a set of outcomes over the post-treatment years, up to 2015 (Table 10). Figure 3 displays the density function of the continuous forcing variable for the full sample and the two subsamples (firms targeted by ML and firms not targeted by ML, respectively). The eligibility cutoff is set at zero: firms to the right of the cutoff are eligible, while firms to the left are not. In order to check whether possible manipulation of the assignment variable is at work, we test the continuity of the density functions at the cutoff for the full sample and the two sub-samples. We employ the test recently proposed by Cattaneo et al. (2017, 2018). As reported in Figure 3, we do not reject the null hypothesis of continuity in all cases. As expected (and necessary for identification), the probability of treatment jumps at the cutoff (Figure 4).

We consider several outcome variables: granted bank loans, sales, investments, probability of adjusted bad loans. We also include both the probability of being credit constrained and the probability of being creditworthy, which we use to test the balancing properties of our sample. Since our sample includes firms that received the GF in 2012 and 2013, outcome variables such as granted bank loans, investments and sales are expressed in terms of their average growth rate in the period 2012-15 or 2013-15 according to the year of treatment. The same averages are computed for non-treated firms, with the initial year being randomly assigned and the proportion of 2012s being the same as that of the treated firms. "Adjusted bad loans" is, instead, a dummy equal to 1 if the firm has adjusted bad loans in 2015 and 0 otherwise.

To substantiate the assumption of randomization in a neighborhood around the eligibility cutoff, on which the fuzzy-RDD strategy is grounded, we perform a series of balancing tests using a set of covariates measured in pre-treatment years. The results, reported in Table 11, show that firms at both sides of the cutoff are characterized by the same level of bank loans, growth rate of bank loans, sales, investments, probability of being credit constrained and probability of being creditworthy. The balancing properties hold for the full sample and, separately, for both of the subsamples according to the value of the ML targeting rule.

Non-parametric estimates of the impact of the GF are reported in Table 12. The first column displays the results for the full sample, while the second and the third report the estimates related to the sample of ML-targeted firms and ML-non-targeted firms, respectively. As this exercise aims at testing the potential heterogeneous effects of the ML targeting rule that we propose, we do not elaborate on the full sample estimates.[15] We therefore focus on the results of columns 2 and 3, which display the fuzzy-RDD estimates carried out separately in the two samples of firms, split according to our ML-targeting rule. In line with our previous descriptive

[15] Although being estimated over a different period with respect to de Blasio et al. (2018), the results in the first column are in line with their findings about the positive impact of the GF on bank loans' growth rate. Moreover, like previous evidence, no impact is found on real outcomes such as sales and investments. Concerning firm riskiness, this more recent evidence suggests that the GF does not lead to an increase in adjusted bad loans. This could partly reflect the different economic trends when the policy was evaluated (de Blasio et al., 2018, focused on the peak of the financial and economic crisis), the change in the rules of the GF introduced in mid-2010 (aiming at improving the screening of firm creditworthiness, among other things), as well as a more selective assessment undertaken by the Fund, following the increase in the amount of the guarantee called on during the early years of the crisis.

findings, the impact of GF on bank loans is positive and statistically significant for the sample of ML-targeted firms, while no effect at all is detected in the sample of ML-non-targeted firms. As for the full sample, the fuzzy-RDD estimates in both the subsamples show no impact of the GF on both sales and investments. Nonetheless, some heterogeneity seems to be at work in both cases, as the sign of the estimated impact is positive for ML-targeted firms and negative for non-ML-targeted ones. This evidence is consistent with the expected real outcomes when the credit guarantee reaches truly credit-constrained firms. Finally, no impact of the GF is detected on adjusted bad loans, in both subsamples.

One issue pertains to the fact that the difference in the effect between the two subsamples might be due to a specific feature among the observables that are used for prediction, rather than to the ML predicted status *per se*. In terms of evaluating the performance of ML targeting, however, it does not matter which is the feature that drives the results. What we want to know is whether using ML targeting in the way we propose identifies, on the basis of a large set of covariates, a group that has larger policy impact. Another potential issue is that compliance might be different in the two subsamples. The RDD identifies the effect for compliers around the threshold, but compliers might be completely different in the two subsamples. However, we also find evidence of heterogeneity in the eligibility effect on granted credit around the threshold, measured by the so-called Intention To Treat (ITT). The latter is, indeed, statistically significant (and slightly larger) for ML targeted firms only. As what matters for us is eligibility, these results corroborated our findings.

6. Pitfalls and implementation issues

In this Section we discuss the problem of contamination due to the fact that our models are trained in a sample where a fraction of the firms is already receiving the guarantee (Subsection 6.1). We then compare the different models in terms of transparency, which might be an important requisite for the policymaker (Subsection 6.2). We also outline the pros and cons of an alternative targeting strategy (Subsection 6.3). Finally, we discuss the fact that different assignment rules might also end up prioritizing some specific categories of firms and generate omitted-payoffs (Subsection 6.4).

6.1 Prediction bias when the policy is already in place

One issue pertaining to the comparison with the actual GF rule is that we estimated and validated the ML models on years during which the Fund was already operational. For this reason the dataset is "contaminated" as it also contains firms that already receive the Fund guarantee. Our actual aim is to predict the credit-constrained and creditworthy conditions in the counterfactual scenario without the guarantee (we define both of them as a binary variable S_0), but for some of the firms we actually observe these conditions in the scenario with the guarantee (S_1). If the guarantee has an impact on constraints and default rates, then S_1 and S_0 are different. Our algorithm has been trained and evaluated to predict the observed status, which is a combination of the two counterfactuals, because what we observe is $S_0(1 - GF) + S_1 GF$, where

$GF = 1[guarantee]$. In general, this implies that the predictive power with respect to S_0 is lower.

If we use the ML predicted probabilities to order firms' applications from the least likely to be a target to the most likely, and select only a fraction of them among the most likely ones, then contamination is a problem only if the guarantee changes the position of firms in the distribution of predicted probabilities. Ideally, we would like to order applications according to $Pr(S_0|X)$, where X are the characteristics used for prediction. So long as looking at $Pr(S_0(1 - GF) + S_1 GF|X)$ leads to the same ordering, contamination does not impact the decision rule based on ML targeting. For instance, there might be no change in the ordering if the guarantee flows to groups, as defined by X, that are more credit constrained, but removes only partially the credit constraints, leaving them on average more rationed than the other groups. At the extreme end of that spectrum, if the guarantee removes all credit constraints in the population of firms, then our ML algorithm would not predict any difference between firms, hence it would not do any better than the current GF eligibility score. However, it might also be the case that the guarantee is strongly concentrated in a group that, without it, would have very large credit constraints (i.e. it has a high $Pr(S_0)$). The impact of the guarantee on this group might be such that firms belonging to it end up having little credit constraints in the observed situation, and therefore they are re-ordered among those predicted to have the lowest likelihood of being credit constrained. In that scenario, a decision rule based on the ML algorithm might therefore say that they should not receive the guarantee, while in fact this is a credit-constrained group with a potential large policy impact.

This issue boils down to the question of gains from ML targeting, as discussed in Subsections 5.1 and 5.2. If the contamination problems annihilates the predictive power of the algorithm, or even makes it worse (with respect to S_0) than a random classifier by excluding relevant groups, then we would find that (i) there is no gain in using it to contract the eligible firms or re-assign the guarantee to some that were excluded; (ii) the causal effect of the guarantee is not larger (or even smaller) in the group that has previously received it. Our results show the contrary and, therefore, ease the concerns about the impact of contamination. Obviously, ML targeting is not perfect as it might still include groups that should not receive the guarantee and exclude others that should, but our aim is to improve over the existent eligibility rule, not to find the perfect "oracle" one. Furthermore, if by "oracle" we mean a rule that chooses the firms with larger causal effect, one should remember that we can never identify each single firm's causal effect, but at most average effects within specific groups.

6.2 Transparency and manipulability

In principle, each prediction model can be used to assess whether a single firm is ML target or not, on the basis of the characteristics available at the time of the application for the GF. However, the models differ in terms of transparency. Our favorite ML prediction model (random forest) is more of a black box, as it does not provide an easily interpretable decision rule. Being an average across a large set of estimated decision trees, the prediction cannot be interpreted by

simply looking at thresholds across different variables. This might be a concern for a policymaker that would tend to favor transparency, and possibly lead firms to raise issues of discrimination as they cannot easily understand why they have been excluded (Athey, 2017). The prediction provided by the decision tree is, instead, the most transparent one, as we select the ML-targeted firms by looking at relatively few variables and comparing them to specific thresholds. This could be more easily communicated as it resembles most of the ordinary policy allocation rules. The LASSO model should be more interpretable, but the presence of interactions makes it less so. Furthermore, the final prediction for LASSO depends on a linear index of a large set of covariates (see the Appendix), and therefore it is not simple to evaluate which characteristics determine whether a firm is eligible or not.

The main trade-off in choosing a simpler algorithm is in terms of accuracy. As already mentioned in Subsection 4.3, random forest performs better than the other methods, decision tree included, in out-of-sample prediction. This is particularly true for the creditworthy status. Furthermore, if instead of looking at the misclassification rate we look at the ranking of predicted probability, the decision tree does particularly badly if we want to select only the groups with highest predicted probability of being credit constrained (or creditworthy). Figures 5a and 5b show the fraction that actually belongs to a status if we were to select as targets only those whose predicted probability of belonging to the status is in the top x fraction. It can be noticed that the decision tree, being overly simple, is not able to discriminate among the highest predicted probability. To decide which methods to implement, a decision maker should trade-off transparency with the amount of misclassification that arises with the chosen rule.

Another possible advantage of a simpler method is the amount of information required. The decision tree, as highlighted in Figures A11 and A12, would require relatively few variables. On the other hand, the random forest model requires a large set of covariates. This is potentially costly, although it should not be forgotten that all the information we use is digitalized in administrative databases and that the GF administrators and, more in general, policymakers, have access to all the information that we use for our predictions.

The trade-off between accuracy on the one hand, and the transparency and information burden on the other, might therefore lead a policymaker to choose a simpler model. Nevertheless, for our application we aim at improving the existing eligibility criteria. With respect to the random forest model, the current GF scoring is possibly easier to assess, as it is based on few budget indices and thresholds that can be summarized in an Excel spreadsheet. In terms of formal transparency, therefore, the current GF rule is more interpretable. However, there is another dimension of transparency, which we can call substantive transparency. This dimension concerns the accountability of the policy maker for accomplishing her mission, which is using public money in an effective way. In this respect, our random forest algorithm might be preferable, because the gains in terms of effectiveness associated with ML targeting are substantial.

Finally, if we look at actual implementation we should not neglect that a simple rule allows each firm and bank to assess eligibility before formally applying for the guarantee. This

facilitates the process, but hinders the ability of the GF board to assess how many firms would be interested but are not eligible. This information might be particularly useful for ex-post evaluation of the GF performance. A more complicated rule, which requires an assessment made through an online platform (after registration) or by the GF itself (after application), may allow a perhaps independent evaluator to focus only on interested firms and use as a control group those that were excluded because they were not eligible.

Another important implementation issue concerns manipulability. Ex-post, when the rule has been defined and made public, applicant firms might alter their variables in order to access the guarantee. This can be done at different levels. The first is to misreport information in the application, but as we use data recorded in digitalized archives we believe this is a minor risk. The second is to alter the variables reported in the archive. This, however, involves fraud and implies a strong legal risk for the applicant. The third is to make some (possibly costly) financial adjustments aimed at meeting the eligibility criteria. We believe this is possible, but this risk is equally shared by the current rule, over which we aim to improve. As the random forest based eligibility rule we propose is even more of a black box, we find it hard for an applicant to carry out this operation. Manipulability can also be an issue ex-ante, where firms behave strategically to alter the variables that we use as proxies for the credit-constrained and creditworthy status. This seems more relevant with respect to the preliminary information system, where requests for access to the system (that we use as a proxy for credit requests) may be performed to alter the dataset and therefore the estimated algorithms. However, individual firms have hardly any chance of influencing the estimates by filling out a loan application (which, in turn, might lead to a request for preliminary information by the bank) when the loan is not needed. Each request counts as one (in a very large dataset) and we also aggregate multiple requests in the same quarter. Furthermore, only credit-constrained firms could be potentially interested in manipulating the algorithm because accessing the guarantee would provide them with benefits by improving their likelihood of obtaining a loan. In this case, there would be no issue of manipulated preliminary information requests, because these "fake" applications would be made by firms that are indeed credit constrained. Finally, presenting a fake loan application would require that a bank receives such a request and, eventually, processes it. It might, therefore, be quite costly to do so, at least in terms of reputational concerns.

6.3 Targeting after the ex-post evaluation

An alternative strategy could be using ML to identify, within an ex-post evaluation framework, the firm's characteristics associated with a stronger payoff. For instance, Ascarza (2018) uses a decision-tree based algorithm to identify which customers should be targeted by firms' retention programs aimed at avoiding churning. The author exploits a pilot in which the treatment (the retention program) was randomly assigned among customers.

One disadvantage of this approach is that it requires an ex-post evaluation setting where it is feasible to neatly study the heterogeneity of the effects across a large set of subgroups, such as a randomized experiment. Secondly, our strategy is grounded in the theories that outline the

features of the firms that ideally should be eligible for guarantee funds in order to reach both additionality and financial sustainability (World Bank, 2015). Hence, our strategy identifies ex-ante eligible firms starting from two observable and credible measures of firm credit constraints and firm creditworthiness.

Only in the second step we employ ex-post evaluation, based on a RDD exercise, to provide evidence that targeting this ML-based group would indeed improve the performance of the GF. Even though the improvements can be appreciated only for the RDD sub-population of interest, we believe that the opposite strategy (i.e. RDD as a first step) would be less reliable for the following reasons. First, using RDD local estimates as a starting point in the search for heterogeneous effects along a large set of non-pre-specified dimensions could be debatable. Second, using RDD local estimates as a starting point for re-targeting would limit the external validity of a targeting exercise to firms that are somehow similar to the sub-population that the RDD estimates refer to.

More in general, our strategy could also be applied to the case in which a policy has not yet been rolled out and it is costly to delay its introduction, as was the case for the GF during the recession. It is, therefore, interesting to understand the pros and cons of such a strategy. In this case, however, the policy should still leave scope for the ex-post evaluation by also making eligible to the policy some firms that are not ML target (for instance a randomly selected group). To assess the heterogeneity ex-post, we need to be able to identify a treatment effect also within the non-targeted group.

6.4 Omitted payoffs

Out ML targeting rule is trained with the aim of increasing the GF effectiveness in raising bank loan availability and reducing the share of loans that go into default. However, any targeting rule, including the current one, might end up having other effects (omitted payoffs, see Kleinberg et al., 2018), which might or might not be desirable.

Given that the GF fund was strongly advocated as a counter-measure for the recession, we examine two important issues. The first is whether the rule tends to favor or not firms in disadvantaged territories that have been strongly hit by the crisis, mostly in the Southern regions. The second is whether the fund tends to flow to banks with certain characteristics, such as being part of a group and having a variety of funding sources. Table 13 shows the correlation between a set of pre-treatment firm characteristics (main bank belonging to a group, number of lending banks, funding gap of the main bank, firm headquarters in Southern Italy) and, in turn, the GF eligibility rule (dummy equal to 1 if the firm is eligible) and the ML targeting rule (dummy equal to 1 if the firm is targeted by ML).

GF eligibility tends to have a bias in favor of firms whose main reference bank (in terms of granted credit) belongs to a group or in favor of firms that have already taken out loans with several banks. Conversely, our favorite ML eligibility is negatively correlated with the firm being more indebted towards a bank that belongs to a group, and tends to prioritize those with

few lending relationships. Moreover, ML targeting seems to favor firms whose main bank has a lower funding gap, i.e. its funding source mainly consists of households deposits. In terms of regional differences, GF eligibility is negatively correlated with the firm being located in the South of Italy, where firms generally face more difficulties in accessing credit, while the opposite holds for ML eligibility. The former, therefore, seems to favor more developed areas. These correlations illustrate that, despite being focused on specific issues, each targeting rule might end up prioritizing firms with certain characteristics. This might satisfy additional (omitted) payoffs that the policymaker has in mind, or even work against them.

7. Conclusions

Gains from ML targeting seem to be relevant. Using the current GF selection mechanism, around 47 per cent of the guarantees (approximately €1,2 billion) went to firms that are not ML targets and showed smaller benefits in terms of access to credit. By using the same amount of public resources, ML could have improved the effectiveness of the GF by replacing about 40 per cent of the current GF beneficiaries with credit constrained and creditworthy non-beneficiary firms, leading to both an increase in the volume of bank loans and in the number of guaranteed firms. We have shown that ML algorithms also come with some downsides in terms of transparency and administrative burden. The current rule might seem formally less opaque, but it fails to be accountable with regard to explaining how it was designed and whether it meets the policy goal of facilitating access to credit for firms that are financially sound but credit constrained. Hence, it is not clear whether we would lose transparency by using, instead, an ML algorithm trained on data and fully evaluated.

All in all, we are aware that the actual implementation of an eligibility mechanism based on ML might prove difficult, also owing to concerns about manipulability and legal constraints that prevent the GF from discriminating on the basis of estimated decision rules. Furthermore, the policymaker might not want to fully apply the ML targeting rule, in order to leave scope for ex-post evaluation of the heterogeneity of effects.[16] Nevertheless, our proposal could be used as a benchmark to compare other rules that satisfy additional requirements imposed by the policymaker. It is also worth stressing that our ML-based eligibility mechanism need not be strictly intended as an alternative to the current scheme. In particular, the two mechanisms could also be used jointly, with the ML-based one arguably used to flag those applications which require a more careful screening (for instance, when the two predicted probabilities of being credit constrained and creditworthy are approaching one, while the baseline selection rule would reject the firm). Finally, we neither provide a full evaluation of the role of the GF in the credit markets nor do we aim to defend its existence. Our purpose is solely to highlight that a better

[16] This can be done, for instance, by randomly excluding a small fraction of ML-eligible firms and randomly including a small fraction of the others. In this way, within the two groups (ML-eligible and non-ML-eligible) there is randomized variation in eligibility and one could assess whether the effect is indeed greater for the ML-eligible group. This modified eligibility rule would essentially state that the probability of being eligible is slightly lower than one for ML-eligible firms and slightly larger than zero for the others. Obviously, as for any other policy, this requires the policymaker and the legal environment to allow for randomization of policies.

targeting via ML could improve a guarantee scheme's effectiveness. How this interacts at the macro level exceeds our aims and requires a different analysis. On a more practical level, it is worth noting that our ML-based targeting algorithm is not a *one-size-fits-all* solution. In particular, our ML-based rule was devised using credit applications and bank loans data recorded in 2011, a period where firms' difficulties in accessing bank credit was at unprecedented levels. We propose an application of our ML-model to policy decisions to be carried out in 2012, where the credit market condition remained largely comparable. Arguably, such a targeting rule, trained within a period of crisis, might prove unfit to guide the policymaker when credit market conditions are liquid and borrower friendly. In general, our ML-based approach would require a constant updating of the targeting algorithm in order to be used for policy purposes.

While our prediction exercise was framed within the GF operations, it has a more general relevance. The prediction of creditworthy firms is also important for private banks. Credit scoring models are already often based on ML algorithms but, since these models are proprietary, we are not in a position to compare our predictions to those. Our ML algorithms might also be useful for supervisory purposes, to double check the accuracy of private forecasts. The prediction of credit-constrained firms is probably even more important from the point of view of aggregate welfare. Knowing who the creditworthy but constrained firms are is important for designing the public interventions justified by credit-market failures. For instance, an important share of the European Union public funds (structural funds) is channeled to lagging regions on the assumption that firms located there have limited access to credit facilities. Our ML targeting might be useful to substantiate this assumption.

References

Abadie, A., Drukker, D., Herr, J. L., and Imbens, G. W. (2004), "Implementing matching estimators for average treatment effects in Stata". The Stata Journal, 4(3): 290-311.

Albertazzi, U., Bottero, M., and Sene, G. (2017), "Information externalities in the credit market and the spell of credit rationing". Journal of Financial Intermediation, 30(-): 61-70.

Andini, M., Ciani, E., de Blasio, G., D'Ignazio, A., and Salvestrini, V. (2018), "Targeting with machine learning: An application to a tax rebate program in Italy". Journal of Economic Behavior and Organization, 156(-): 86-102.

Ascarza, E. (2018), "Retention futility: Targeting high risk customers might be ineffective". Journal of Marketing Research, 55(1): 80-98.

Athey, S. (2017), "Beyond prediction: Using big data for policy problems". Science, 255(6324): 483-485.

Athey, S. (2018), "The impact of machine learning on economics". Available at: http://www.nber.org/chapters/c14009.pdf.

Barone, G., Mirenda, L., and Mocetti, S. (2016), "Losing my connection: The role of interlocking directorates". Rimini Centre for Economic Analysis Working Papers n. 16/09.

Beck, T., Klapper, L. F., and Mendoza, J. C. (2008), "The typology of partial credit guarantee funds around the world". Journal of Financial Stability, 6(1): 10-25.

Belloni, A., Chernozhukov, V., and Hansen, C. (2014), "High-dimensional methods and inference on structural and treatment effects". Journal of Economic Perspectives, 28(2): 29-50.

Carmignani, A., de Blasio, G., Demma, C., and D'Ignazio, A. (2017), "Urbanization and firm access to credit". Bank of Italy, mimeo.

Cattaneo, M. D., Jansson, M., and Ma, X. (2017), "Simple local polynomial density estimators". Working paper, University of Michigan.

Cattaneo, M. D., Jansson, M., and Ma, X. (2018), "Manipulation testing based on density discontinuity". The Stata Journal, 18(1): 234-261.

Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., and Mullainathan, S. (2016), "Productivity and selection of human capital with machine learning". American Economic Review, 106(5): 124-127.

Comitato di gestione del Fondo di garanzia. Various years. Annual reports.

Debashis, G., and Chinnaiyan, A. M. (2005), "Classification and selection of biomarkers in genomic data using LASSO". Journal of Biomedicine and Biotechnology, 2005(2): 147-154.

de Blasio, G., De Mitri, S., D'Ignazio, A., Finaldi Russo, P., and Stoppani, L. (2018), "Public guarantees to SME borrowing. A RDD evaluation". Journal of Banking and Finance, 96(-): 73-86.

Deelen, L., and Molenaar, K. (2004), "Guarantee funds for small enterprises. A manual for guarantee fund managers". International Labour Organization. Available at: http://www.ilo.org/public/libdoc/ilo/2004/104B09_435_engl.pdf.

European Central Bank (2015), "Survey on the access to finance of enterprises in the euro area". Frankfurt: European Central Bank. Available at: https://www.ecb.europa.eu/pub/pdf/other/SAFE_website_report_2014H2.en.pdf?56935ca239cc0aab853703c9b2103145.

Fan, J., and Lv, J. (2008), "Sure independence screening for ultrahigh dimensional feature space". Journal of the Royal Statistical Society Series B, 70(5): 849-911.

Friedman, J., Hastie, T., and Tibshirani, R. (2010), "Regularization paths for generalized linear models via coordinate descent". Journal of Statistical Software, 33(1): 1-22.

Galardo, M., Lozzi, M., and Mistrulli, P. E. (2017), "Social capital, uncertainty and credit supply: Evidence from the global crisis". Bank of Italy, mimeo.

Glaeser, E. L., Kim, H., and Luca, M. (2017), "Nowcasting the local economy: Using Yelp data to measure economic activity". National Bureau of Economic Research Working Papers n. 24010.

Hastie, T., Tibshirani, R., and Friedman, J. (2009), "The elements of statistical learning: Data mining, inference, and prediction". New York: Springer.

Honohan, P. (2010), "Partial credit guarantees: Principles and practice". Journal of Financial Stability, 6(1): 1-9.

Imbens, G. W., and Kalyanaraman, K. (2012), "Optimal bandwidth choice for the regression discontinuity estimator". Review of Economic Studies, 79(3): 933-959.

Jiménez, G., Ongena, S., Peydrò, J. L., and Saurina, J. (2012), "Credit supply and monetary policy: Identifying the bank balance-sheet channel with loan applications". American Economic Review 102(5): 2301-2326.

Jiménez, G., Ongena, S., Peydrò, J. L., and Saurina, J. (2014), "Hazardous times for monetary policy: What do twenty-three million bank loans say about the effects of monetary policy on credit risk-taking?". Econometrica, 82(2): 463-505.

Kleinberg, J., Ludwig, J., Mullainathan, S., and Obermeyer, Z. (2015), "Prediction policy problems". American Economic Review, 105(5): 491-495.

Kleinberg, J., Lakkaraju, H., Leskovec, J., and Mullainathan, S. (2018), "Human decisions and machine predictions". The Quarterly Journal of Economics, 133(1): 237-293.

Lee, D. S., and Lemieux, T. (2010), "Regression discontinuity designs in economics". Journal of Economic Literature, 48(2): 281-355.

McBride, L., and Nichols, A. (2015), "Improved poverty targeting through machine learning: An application to the USAID poverty assessment tools". Available at: http://www.econthatmatters.com/wp-content/uploads/2015/01/improvedtargeting_21jan2015.pdf.

Ministero dello Sviluppo economico (2015), "Relazione sugli interventi di sostegno alle attività economiche e produttive". Available at: http://www.sviluppoeconomico.gov.it/images/stories/pubblicazioni/Relazione_2015.pdf.

Mullainathan, S., and Spiess, J. (2017), "Machine learning: An applied econometric approach". Journal of Economic Perspectives, 31(2): 87-106.

OECD (2013), "SME and entrepreneurship financing: The role of credit guarantee schemes and mutual guarantee societies in supporting finance for small and medium-sized enterprises", Final Report, January 2013.

Poolsawad, N., Kambhampati, C., and Cleland, J. G. F. (2014), "Balancing class for performance of classification with a clinical dataset". In: Proceedings of the World Congress on Engineering, vol. I. Available at: http://www.iaeng.org/publication/WCE2014/WCE2014_pp237-242.pdf.

Riding, A., Madill, J., and Haines, G. J. (2007), "Incrementality of SME loan guarantees". Small Business Economics, 29(1): 47-61.

Saadani, Y., Zsofia, A., and de Rezende, R. (2011), "A review of credit guarantee schemes in the Middle East and North Africa region". World Bank Policy Research Working Paper n. 5612.

Schivardi, F., Sette, E., and Tabellini, G. (2017), "Credit misallocation during the European financial crisis". Bank of Italy Working Papers n. 1139.

Uesugi, I., Sakai, K., and Yamashiro, G. (2010), "The effectiveness of public credit guarantees in the Japanese loan market". Journal of the Japanese and International Economies, 24(4): 457-480.

Varian, H. R. (2014), "Big data: New tricks for econometrics". Journal of Economic Perspectives, 28(2): 3-28.

Vogel, R. C., and Adams, D. W. (1997), "Costs and benefits of loan guarantee programs". The Financier, 4(1-2): 22-29.

World Bank (2015), "Principles for public credit guarantee schemes for SMEs". Available at: http://documents.worldbank.org/curated/en/576961468197998372/pdf/101769-REVISED-ENGLISH-Principles-CGS-for-SMEs.pdf.

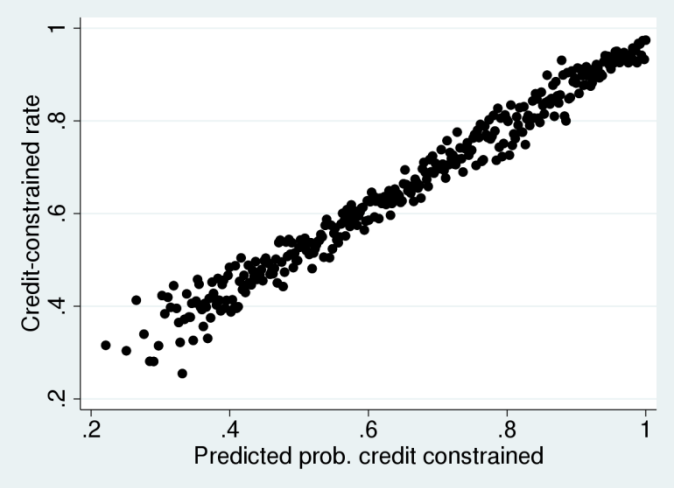
Zhao, Y., and Cen, Y. (2014), "Data mining applications with R". Oxford Academic Press: Elsevier.

Zia, B. H. (2008), "Export incentives, financial constraints, and the (mis)allocation of credit: Micro-level evidence from subsidized export loans". Journal of Financial Economics, 87(2): 498-527.

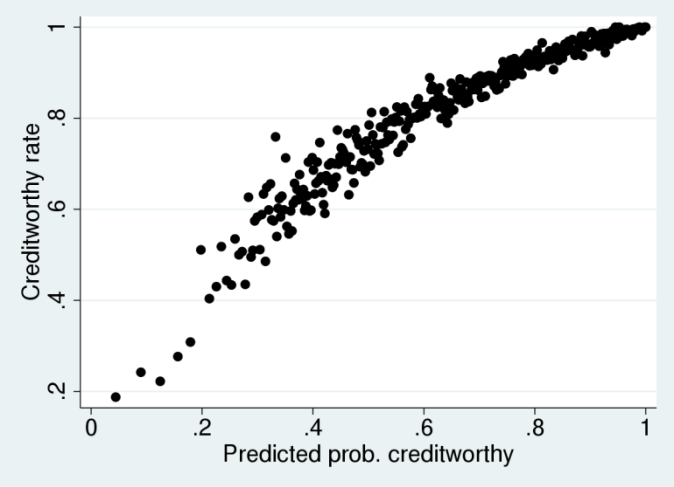
Figures

Figure 1. Random forest predictions

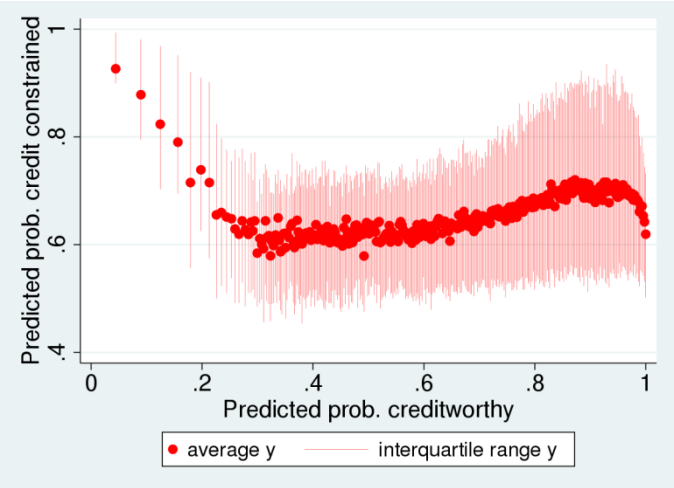
(a) Actual credit-constrained rate vs predicted probability



(b) Actual creditworthy rate vs predicted probability



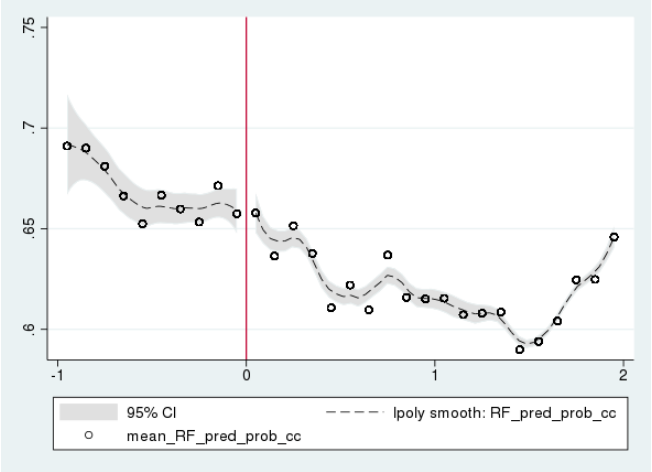
(c) Predicted probability of being credit constrained vs creditworthy



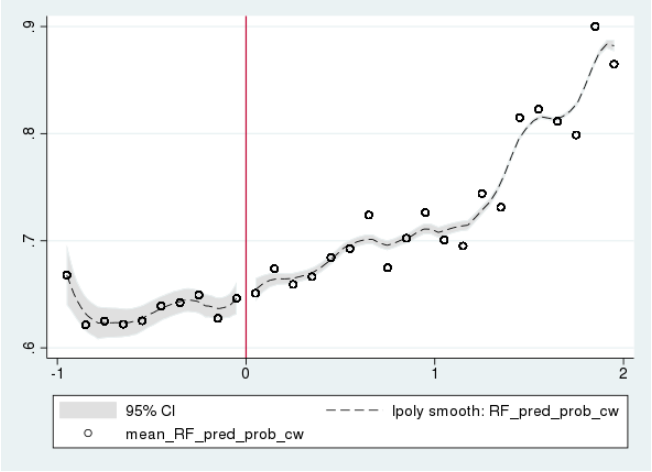
Notes. Testing sample (2011). Each point represents one of 1,000 percentile bins of the variable on the x-axis.

Figure 2. Probability of being an ML target vs actual GF eligibility

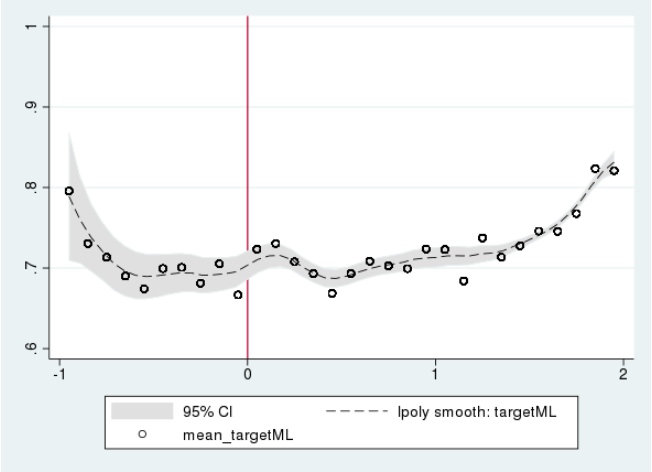
(a) Predicted credit-constrained status vs GF eligibility



(b) Predicted creditworthy status vs GF eligibility



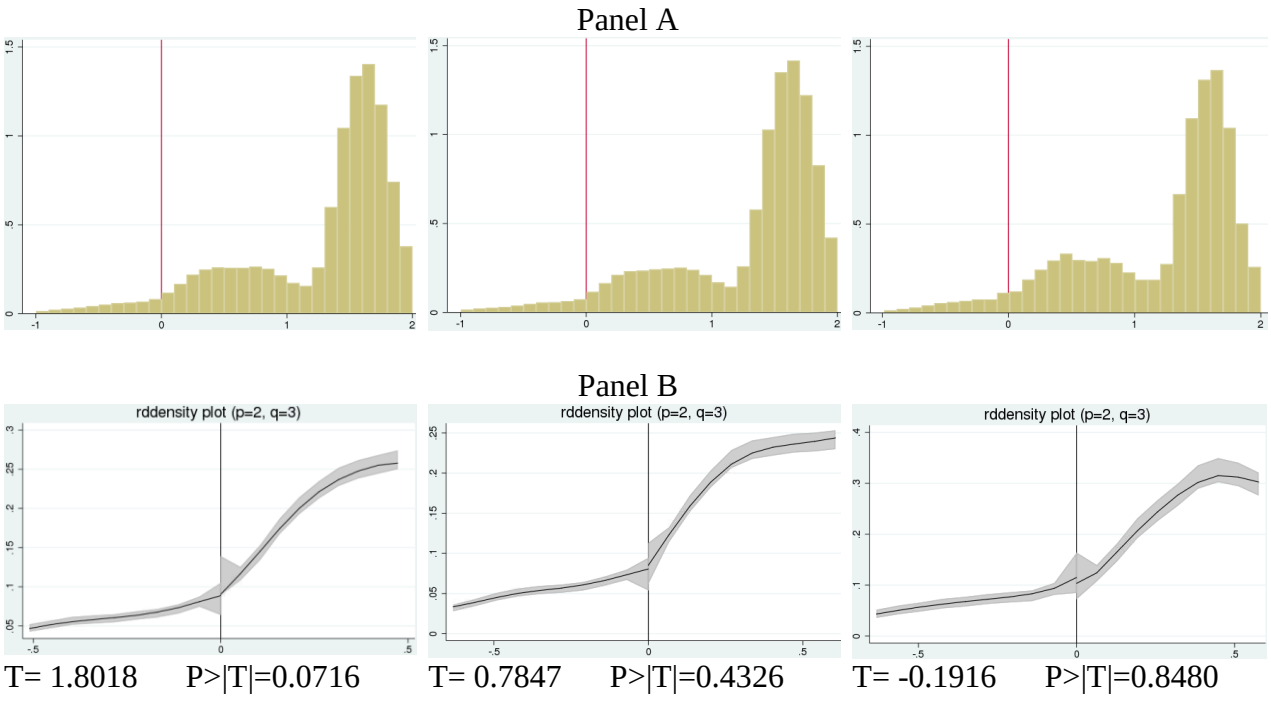
(c) Predicted ML target status vs GF eligibility



Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme (see Subsection 4.4). The x-axis is a continuous measure of GF eligibility (see de Blasio et al., 2018). Eligible firms have $x \geq 0$. The y-axis is the fraction with predicted status equal to 1 (random forest predicted probability ≥ 0.5 in panels a and b and joint predicted status for panel c).

Figure 3. Density function of the forcing variable

(a) Full sample (b) ML target = 1 (c) ML target = 0

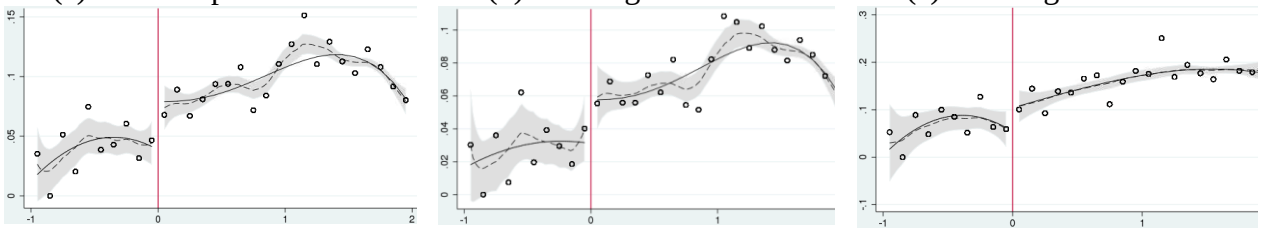


T= 1.8018 P>|T|=0.0716 T= 0.7847 P>|T|=0.4326 T= -0.1916 P>|T|=0.8480

Notes. Selected sample of 62,994 firms (see Subsection 5.2). Panel A: density function of the forcing variable (a continuous measure of GF eligibility; eligible firms have $x \geq 0$), with the eligibility threshold set at 0. Panel B: manipulation tests using local polynomial density estimation (Cattaneo et al., 2017, 2018). $H_0: \lim_{x \uparrow \bar{x}} f(x) = \lim_{x \uparrow \bar{x}} (\bar{x})$. Under the appropriate assumptions, the test statistic T is distributed as a N(0,1). For each indicator, plots of the manipulation test (above) and test statistics (below) are provided.

Figure 4. Probability of treatment at the cutoff

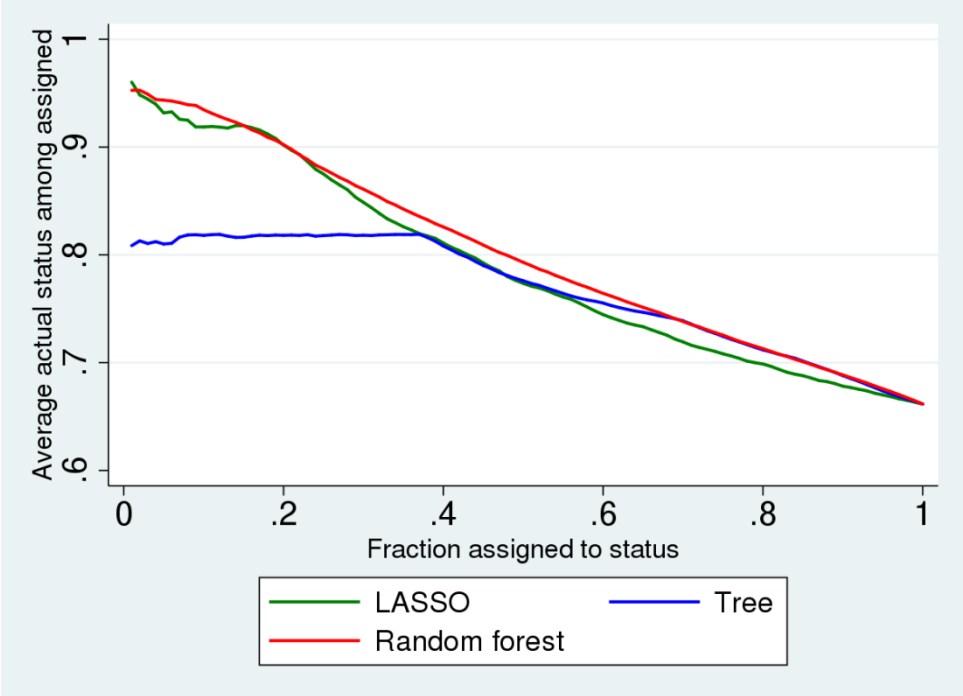
(a) Full sample (b) ML target = 1 (c) ML target = 0



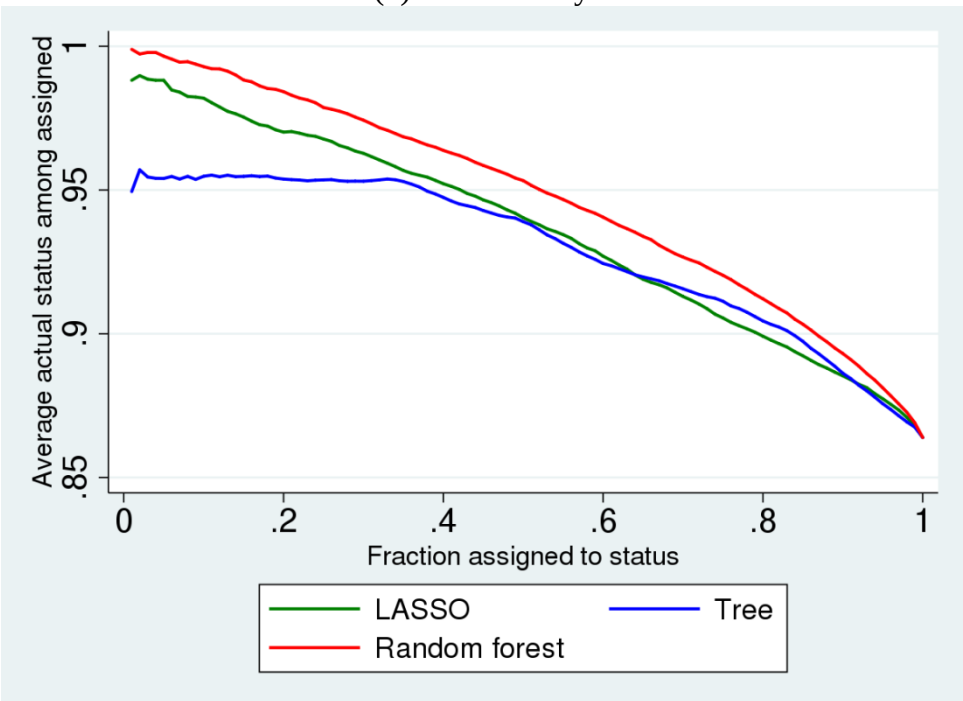
Notes. Selected sample of 62,994 firms (see Subsection 5.2). The x-axis id the forcing variable (a continuous measure of GF eligibility; eligible firms have $x \geq 0$). The y-axis is the fraction of firms that are treated (i.e. GF beneficiary).

Figure 5. Fraction in the actual status in the x-fraction with highest predicted probability

(a) Credit constrained



(b) Creditworthy



Notes. Testing sample (2011). On the horizontal axis the percentage of observations classified as positive, choosing first those with the highest predicted probability (and, therefore, assigned to status). In cases in which multiple observations have the same predicted probability, we chose among them randomly. On the vertical axis the fraction of true positive cases over those classified as positive.

Tables

Table 1. Confusion matrix for the final target

	Y_{pred} = Not Target	Y_{pred} = Target	Misclassification rate: 36.76%	
Y_{actual} = Not Target	17,822	22,779	TN: 43.89%	FN: 21.81%
Y_{actual} = Target	11,451	41,047	FP: 56.1%	TP: 78.18%

Notes. Testing sample (2011). Y_{actual} is 1 if the actual status is to be credit constrained and creditworthy, 0 otherwise; Y_{pred} is 1 if a credit-constrained and creditworthy observation is predicted (predicted probability of each status ≥ 0.5), 0 otherwise. FP is the *false positive rate* computed as the percentage of observations predicted positive, but that are actually negative, over the total number of actually negative observations; TP is the *true positive rate* computed as the percentage of observations predicted positive, that are actually positive, over the total number of actually positive observations; FN is the *false negative rate* computed as the percentage of observations predicted negative, but that are actually true, over the total number of actually positive observations; TN is the *true negative rate* computed as the percentage of observations predicted negative, but that are actually negative, over the total number of actually negative observations.

Table 2. Replication of the GF screening mechanism

	GF eligible (B)		
GF beneficiary (A)	0	1	Total
0	4,518	77,073	81,591
1	160	6,751	6,911
Total	4,678	83,824	88,502

Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme. (A): firms that received (=1) or did not receive (=0) the Fund guarantee over the period 2012-13. (B): firms that are eligible (=1) or not eligible (=0) according to the actual Fund eligibility scoring mechanism, with the scoring procedure based on firm balance sheet data from Cerved group.

Table 3. GF eligibility vs ML targeting

	ML target (B)		
GF eligible (A)	0	1	Total
0	1,174	3,504	4,678
1	16,860	66,964	83,824
Total	18,034	70,468	88,502

Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme. (A): firms that are eligible (=1) or not (=0) for the Fund guarantee according to the actual GF scoring mechanism. (B): firms that are selected as target (=1) or not (=0) by the ML algorithm (random forest).

Table 4. Characteristics of the GF-eligible firms that are not targeted by ML

	Credit constrained (B)		
Creditworthy (A)	0	1	Total
0	874	4,395	5,269
1	11,591	0	11,591
Total	12,465	4,395	16,860

Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme; subset of firms that are eligible according to the Fund rules but that are not targeted by the ML algorithm. (A): firms predicted as creditworthy (=1) or not (=0) by the ML algorithm (random forest). (B): firms predicted as constrained (=1) or not (=0) by the ML algorithm (random forest).

Table 5. GF eligibility and ML predicted firm characteristics

Dependent variable: eligibility for the Fund	Coef.
ML predicted probability of being creditworthy	0.2506003*** (0.0044215)
ML predicted probability of being credit constrained	-0.1186898*** (0.0044105)
Manufacturing, construction, fishing and tradable sector	0.0166327*** (0.0015128)
Constant	0.8238488*** (0.0045158)
Observations	88,502
Adj R-squared	0.0407

Notes. *** p-val $\leq$ 0.01. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme. Linear probability model. The dependent variable is binary, taking value=1 if firms are eligible for the guarantee according to the Fund's rules, and zero otherwise. Standard errors in parentheses. The predicted probabilities refer to the random forest model.

Table 6. ML targeted vs beneficiary firms

	ML target (B)		
GF beneficiary (A)	0	1	Total
0	15,165	66,426	81,591
1	2,869	4,042	6,911
Total	18,034	70,468	88,502

Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme. (A): firms that obtained the Fund guarantee in the period 2012-13. (B): firms predicted as target (=1) or not (=0) by the ML algorithm (random forest).

Table 7. Observed performance of treated firms according to the ML targeting algorithm (contraction)

	ML target=1	ML target=0	Difference	p-value	t-stat
Granted loans 2011-15	0.08	-0.48	0.56	0	6.41
Number of banks 2011-15	0.21	0.08	0.13	0.03	2.15
Adjusted bad loans 2015	0.03	0.06	-0.03	0	5.74
Fixed assets 2011-15	-0.03	0	-0.03	0.25	1.15
Sales 2011-15	-0.17	-0.24	0.07	0.01	2.65
Roa 2015	0.01	0	0.01	0.31	1.01
Gross oper. margin on assets 2015	0.06	0.04	0.03	0.01	2.57
Observations	4,042	2,869			

Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme that received the guarantee. ML target=1: average performance in the period 2011-15 for the set of firms that received the Fund guarantee and are targeted by our ML algorithm (combined prediction from random forest). ML target=0: average performance in the period 2011-15 for the set of firms that received the Fund guarantee and are not targeted by our ML algorithm. Adjusted bad loans, Roa and Gross operating margin on assets are ratios measured in 2015. Number of banks is a growth rate computed as $[var(2015)/var(2011)] - 1$. The other variables are growth rates are computed as: $\log[var(2015) + 1] - \log[var(2011) + 1]$ with the original var measure in euro

Table 8. Re-ranking

	Re-ranking=1	Re-ranking=0	Difference	p-value	t-stat
Granted loans 2011-15	0.18	-0.48	0.66	0	6.03
Number of banks 2011-15	0.22	0.08	0.14	0.13	1.53
Adjusted bad loans 2015	0.03	0.06	-0.02	0	3.33
Fixed assets 2011-15	-0.08	0	-0.08	0.05	1.98
Sales 2011-15	-0.16	-0.24	0.07	0.05	1.97
Roa 2015	0.02	0	0.02	0.22	1.22
Gross oper. margin on assets 2015	0.06	0.04	0.03	0.11	1.58
Observations	1,303	2,869			

Notes. 2012 sample, only firms belonging to the sectors that are currently eligible for the GF scheme. Re-ranking=1: theoretical values of the average performance in the period 2011-15 of the subset of non-eligible, non-treated firms that were targeted by our ML algorithm. The theoretical values are computed by means of a matching procedure that associates to each of these firms (by means of nearest neighbor matching) a firm that was targeted by ML and that received the Fund guarantee. In particular, we consider 3,504 firms that are targeted by ML but not eligible according to the Fund rules (Table 3); for 1,303 among them we manage to find a match firm belonging to the group of those that received the Fund guarantee and were targeted by ML. The column reports the average performance of the matched-treated firms. Re-ranking=0: average performance in the period 2011-15 for the subset of firms that received the Fund guarantee and are not targeted by our ML algorithm. Adjusted bad loans, Roa and Gross operating margin on assets are ratios measured in 2015. Number of banks is a growth rate computed as $[var(2015)/var(2011)] - 1$. The other variables are growth rates are computed as: $\log[var(2015) + 1] - \log[var(2011) + 1]$ with the original var measure in euro.

Table 9. Futile expenditure

GF beneficiary	Amount financed (million euro)	Guarantees (million euro)	Firms	Average amount financed (thousand euro)	Average guarantee (thousand euro)
Non-ML target	1,200.3	718.3	2,869	418.4	250.4
of which:					
non-credit constrained, creditworthy	836.5	510.0	1,722	485.7	296.2
credit constrained, non-creditworthy	258.7	147.0	852	303.7	172.5
non-credit constrained, non-creditworthy	105.1	61.3	295	356.4	207.8
ML target	1,335.6	828.0	4,042	330.4	204.9

Notes. 2012 sample, only firms that received the guarantee.

Table 10. Fuzzy-RDD analysis, sample

	(a) Full sample	(b) ML target = 1	(c) ML target = 0
Treated	6,294	3,587	2,707
Not treated	56,700	42,994	13,706
All firms	62,994	46,581	16,413

Notes. Selected sample of 62,994 firms (see Subsection 5.2).

Table 11. Fuzzy-RDD analysis, balancing properties

	(a) Full sample	(b) ML target = 1	(c) ML target = 0
Panel A. Pre-treatment bank granted credit (level)			
Numerator	-0.1684123 (0.1543444)	-0.0856334 (0.208376)	-0.108946 (0.1903549)
Denominator	0.0277192** (0.013287)	0.0238867 (0.0149744)	0.0459006* (0.0273586)
Wald	-6.075649 (6.307667)	-3.584978 (9.055825)	-2.373518 (4.306304)
Panel B. Pre-treatment bank granted credit (growth rate)			
Numerator	0.0244678 (0.0248359)	-0.0129047 (0.03137)	0.1272396** (0.0536416)
Denominator	0.0272204*** (0.0099534)	0.0263203** (0.0118676)	0.0322442 (0.024307)
Wald	0.8988773 (0.9456948)	-0.4902954 (1.227729)	3.946121 (3.304303)
Panel C. Pre-treatment sales (level)			
Numerator	-0.1529719 (0.1058798)	-0.1011885 (0.113627)	-0.2576026* (0.132231)
Denominator	0.03219** (0.0155781)	0.0233493 (0.0147553)	0.0415998 (0.0258571)
Wald	-4.752152 (4.062968)	-4.333679 (5.684743)	-6.192405 (4.931695)
Panel D. Pre-treatment investments (growth rate of fixed assets)			
Numerator	0.1008455** (0.0441946)	0.1443115*** (0.0506535)	0.0160636 (0.0776696)
Denominator	0.0252813** (0.0117551)	0.0244879** (0.0124597)	0.036326 (0.0232882)
Wald	3.988933 (2.581401)	5.893168 (3.656554)	0.4422085 (2.172141)
Panel E. Prob. of being credit constrained			
Numerator	-0.0046135 (0.0058763)	0.0007804 (0.0052879)	-0.0347894*** (0.0123218)
Denominator	0.026046*** (0.0098075)	0.0229268** (0.0096081)	0.0365576* (0.0209524)
Wald	-0.1771297 (0.2293389)	0.03404 (0.2318243)	-0.9516312 (0.6162278)
Panel F. Prob. of being creditworthy			
Numerator	0.0205594** (0.0103005)	-0.0066693 (0.0074768)	0.0287371 (0.019476)
Denominator	0.0298048** (0.0138033)	0.0228396* (0.0118621)	0.0516234* (0.0290012)
Wald	0.6898023 (0.4858961)	-0.2920045 (0.3569744)	0.5566675 (0.4992364)

Notes. *** p-val $\leq$ 0.01, ** p-val $\leq$ 0.05, * p-val $\leq$ 0.1. Selected sample of 62,994 firms (see Subsection 5.2). Fuzzy-RDD non parametric estimates. The optimal bandwidth was retrieved by Imbens and Kalyanaraman (2012) procedure. Outliers below the 5th or above the 95th percentile were dropped. Standard errors in brackets.

Table 12. Fuzzy-RDD analysis, non-parametric estimates

	(a) Full sample	(b) ML target = 1	(c) ML target = 0
Panel A. Bank granted credit (growth rate)			
Numerator	0.016817** (0.0068844)	0.0184388** (0.0077351)	0.0151382 (0.0132754)
Denominator	0.0257794** (0.0099872)	0.0232211** (0.0096631)	0.0349826 (0.0240806)
Wald	0.6523421* (0.3633817)	0.7940563* (0.4681124)	0.4327344 (0.4666691)
Panel B. Sales (growth rate)			
Numerator	0.0069459 (0.0107999)	0.0180564 (0.0122835)	-0.016244 (0.0188225)
Denominator	0.0260874*** (0.0092224)	0.0198545** (0.0089386)	0.0384457* (0.0202947)
Wald	0.2662543 (0.4242587)	0.9094354 (0.7420002)	-0.4225176 (0.5382537)
Panel C. Investments (growth rate of fixed assets)			
Numerator	-0.0001976 (0.0073003)	0.0115078 (0.0095961)	-0.033853** (0.015784)
Denominator	0.0295791*** (0.0092208)	0.0232902** (0.0108172)	0.0421291* (0.0234695)
Wald	-0.0066788 (0.2467698)	0.4941046 (0.4911585)	-0.8035543 (0.5997158)
Panel D. Prob. of adjusted bad loans			
Numerator	-0.002043 (0.0113976)	-0.0000369 (0.0132788)	-0.0051727 (0.0238178)
Denominator	0.0228931** (0.0092405)	0.0176069* (0.0106632)	0.0371323* (0.0197966)
Wald	-0.0892417 (0.5001871)	-0.0020936 (0.7542111)	-0.1393052 (0.646751)

Notes. *** p-val $\leq$ 0.01, ** p-val $\leq$ 0.05, * p-val $\leq$ 0.1. Selected sample of 62,994 firms (see Subsection 5.2). Fuzzy-RDD non parametric estimates. The optimal bandwidth has been retrieved by Imbens and Kalyanaraman (2012) procedure. Outliers below the 5th or above the 95th percentile were dropped. Standard errors in brackets.

Table 13. GF and ML eligibility and omitted payoffs

	Y = main bank belongs to a group (pre-treatment)		Y = number of funding banks (pre-treatment)		Y = funding gap of the main bank (pre-treatment)		Y = firm headquartered in Southern Italy	
GF eligible (1)	0.0432***		0.451***		0.906		-0.0245***	
	(0.00737)		(0.0401)		(0.640)		(0.00620)	
ML Target (2)		-0.148***		-2.608***		-1.571***		0.0377***
		(0.00367)		(0.0309)		(0.265)		(0.00316)
Manuf. & constr. sectors	0.122***	0.110***	1.327***	1.112***	0.532**	0.454*	-0.0211***	-0.0183***
	(0.00320)	(0.00319)	(0.0221)	(0.0210)	(0.242)	(0.242)	(0.00267)	(0.00269)
Constant	0.546***	0.709***	1.884***	4.480***	14.13***	16.22***	0.227***	0.172***
	(0.00727)	(0.00360)	(0.0390)	(0.0303)	(0.645)	(0.258)	(0.00613)	(0.00312)
Observations	88,502	88,502	88,502	88,502	63,299	63,299	88,502	88,502

Notes. *** p-val ≤ 0.01, ** p-val ≤ 0.05, * p-val ≤ 0.1. Robust standard errors in brackets. (1) firms that are eligible for the GF; (2) firms that are targeted by ML. For data availability reasons, the funding gap of the main bank can be computed only for a subset of observations.

Appendix

This Appendix provides additional information on a number of topics: the dataset and its peculiarities (Sections A1-A2); the strategy we follow for model selection and training (Section A3); details on the implementation of the ML algorithms to predict credit-constrained and creditworthy firms (Sections A4-A5, which are meant for readers interested in more technical elements); a comparative description of model prediction results as well as model selection (Section A6).

A1. Covariates description and data cleaning

Our main data sources are: Credit register (CR) data on firms' credit history and bad loans; Cerved data on firms' balance sheets. From CR we extract quarterly data covering the two years preceding the quarter when the firm issues the loan request. In particular, we consider: (i) the amount of total bank loans granted to the firms; (ii) the amount of total bank loans granted and actually used by the firm; (iii) the total number of banks' lending to the firm; and (iv) a dummy variable indicating whether the firm has been reported as having bad loans. In addition, we include (v) a binary variable to identify firms about which we have no credit history data within the CR dataset (most of these firms likely have lending relationships with some banking institutions, but they do not show up in the data because the total amount of loans granted by each institution does not reach the €30,000 CR threshold). As for balance sheet data, we select from the Cerved database the two most recent annual observations available before the PI request (loan application). In particular, we consider a set of balance-sheet items, taken from both balance sheets and income statements. We also include some indicators such as the return on assets, operating margin on assets and the leverage index. In addition, we generate a dummy variable identifying firms with negative or null equity. The list of covariates is reported in Table A1, while descriptive statistics can be found in Table A2.

After a data cleaning procedure designed to remove missing data, we try to reduce the information redundancy by analyzing the pairwise correlation among the covariates. Since we are dealing with both categorical and numerical variables, we rely on three different correlation statistics: the Pearson correlation index, the Polyserial correlation index and the Tetrachoric correlation index.[17] Using these statistics, we: (i) select some variables among the ones too highly correlated (more than 95 per cent); (ii) discard variables that are almost not correlated with the dependent variable (a correlation coefficient smaller than 5 per cent). The variables that pass the screening procedure are reported in Tables A3 and A4.

[17] The Pearson correlation index measures linear correlation between two numerical variables. The Polyserial correlation index is an index of bivariate association among numerical and categorical variables, resulting from an underlying continuous variable. The Tetrachoric correlation index measures the agreement for binary data. It estimates what the correlation would be if measured on a continuous scale.

A2. Some peculiarities of the 2011 and 2012 datasets

In the 2011 dataset the unit of observation is the loan application of a given firm in a given quarter of the year. The number of observations in the sample as well as the number of corresponding firms and quarters is reported in the main text. The same firm might appear more than once (up to 4 times) within the dataset. As a consequence, the same firm can be observed both as credit constrained and not constrained, depending on the quarter when the PI is issued. This is due to the fact that a firm is defined as credit constrained according to the dynamics of its bank loans in the six months following the PI request. Hence, PI issued in different quarters are associated with different time windows. Concerning the observed status of creditworthy, instead, there is no such variability, as we only consider whether firms have adjusted bad loans or not at the end of 2014.

A similar pattern is observed for ML-based predictions. In particular, the same firm can be ML-predicted as credit constrained or not, or creditworthy or not, depending on the quarter when the PI was issued. In particular, if a firm has a PI issued before June, then the ML algorithms will use firm balance sheet data at t-2, while if the firm has a PI issued after June the ML algorithms will use data at t-1, because balance sheet data at t-1 are usually made available in June. For instance, consider a firm that has two PIs, one in May 2011 and one in July 2011. The firm does not have adjusted bad loans in 2014; hence the observed creditworthy status is 1. When we predict the creditworthy status of this firm, we will be using 2009 balance sheet data in the first case and 2010 balance sheet data in the second. It is possible that in one case the firm will be predicted as not creditworthy and, in the other case, it will be predicted as such.

Unlike the 2011 dataset, the 2012 one is a cross-sectional dataset obtained after a random sample selection of one quarter occurrence for each firm. We apply our ML rule to firms for which banks have issued a PI request in 2012. As in the 2011 dataset, it is possible that the same firm has PI requests in different quarters of the year. However, we use this sample to simulate a policy scenario, where each firm is either a beneficiary or not of the GF, and each firm is either a ML target or not. Since the same firm with a PI request issued in different quarters might be associated with different ML predictions (if they rely on balance sheet data in different years), in those cases we randomly selected only one occurrence and discarded the remaining one(s). This leads to a drop in the number of observations, but not of the firms. This also means that, in the resulting 2012 dataset, observations and firms coincide (differently than the 2011 dataset). A further drop depends on the fact that, in order to replicate the GF eligibility mechanism, we need to gather a large set of balance sheet data from 2009, which are not available for all the firms in our sample. Finally, as we want to compare the GF eligibility mechanism with the ML targeting rule, we also restrict our sample of firms to those who belong to the GF eligible sectors. This leaves us with a sample of about 88,000 firms.

A3. Strategy for model selection and training

For the decision tree, we implement a top-down approach that uses the minimization of Gini impurity as a splitting criterion. The complexity parameter (c_p) for pruning the tree is chosen using 10-fold cross-validation over the interval $[\alpha, \infty)$. The value α is chosen by considering a not too small α so that we do not deal with splits leading to leaves in which the classes frequencies are almost equal. Looking at the cross-validation errors for a grid of possible complexity parameters, we choose the smallest c_p whose associated error is larger than the minimum error achieved in the cross-validation plus its standard deviation. This is done because the error usually reaches a plateau around the c_p which gives the minimum error, and therefore by taking a larger (but close enough) c_p we reduce the risk of over-fitting (by reducing complexity) keeping a similar cross-validated error.

As for the random forest, the input parameters for this algorithm are the number of trees that constitute the forest (n) and the number of variables used to fit each tree (m). We validate these parameters looking at the *out-of-bag* (OOB) misclassification error.[18] We allow the number of variables m to vary from 1 to $\sqrt{P}$ where P is the total number of covariates in the (post-screening) X matrix. We instead allow the maximum number of trees to be such that the probability that each variable does not appear in any tree is very low (approximately 10^{-6}). We then choose the combination (n, m) that has the minimum OOB error.

The third algorithm we use is the LASSO regression in its logit specification. In this framework, one may account for the potential role played by non-linearities by generating all pairwise interactions between the explanatory variables included in the observables set (say X_1 for the credit-constrained exercise and X_2 for the creditworthy one). Since this procedure leads to a marked increase in the dimension of the covariates matrix in each exercise, we apply an additional screening process to select only those that are more correlated with the respective dependent variable (dropping those with correlation coefficient smaller than 5 per cent).[19] We use the two matrices thus obtained to estimate our predictive models, including 32 variables in the first exercise and 71 in the second. In line with Debashis and Chinnaiyan (2005), we choose the penalization parameter (λ) that minimizes the loss function; the optimal λ is selected through cross-validation using the one-standard-error rule.[20]

If there are strongly unbalanced classes in the sample, then this imbalance might bias the ML classifier towards the over-represented class. The classifier might end up having a high misclassification error for the under-represented class. In this circumstance, a rebalancing procedure should be applied. This is the case for the "creditworthy status", where the distribution

[18] The OOB error is computed as follows: for each observation we consider all the trees estimated on bootstrapped training sets where that observation does not appear, and we use their predictions to compute the misclassification error.

[19] After the inclusion of all interaction terms, the set of explanatory variables counts, respectively, 152 units in the credit-constrained exercise and 189 in the creditworthy exercise (in both cases, we exclude interactions that generated uninformative and invariant constant terms).

[20] The loss function to minimize is the cross-validation misclassification rate. The λ parameter is validated over a grid of multiple values within the interval $[\lambda_{min}, \lambda_{max}]$, where λ_{max} is defined as discussed in Friedman et al. (2010).

of the creditworthy vs not creditworthy is strongly unbalanced as the not creditworthy observations are about 14 per cent of the total. Following Poolsawad et al. (2014) and Zhao and Cen (2014), we adopt an under-sampling strategy to solve the class imbalance. In particular, we randomly select only a subset of the observations belonging to the over-represented class and discard the remaining ones, so that the number of majority class observations (creditworthy firms, in our case) in the training set equals twice the number of under-represented class (not creditworthy firms) observations.[21]

While we estimate two separate models, another approach could be to predict directly the target firms as those that are both constrained and creditworthy without using the balancing procedure. With this new dependent variable, we observe two things: (i) the percentage of observations in the constrained status is about 66 per cent of the training set, only 10 percentage points higher than that of the observations in the final target (56 per cent) meaning that, in this case, the balance feature of the target vs non-target status is informed by the credit-constrained status rather than by the creditworthy one; (ii) no improvement is reached in terms of misclassification error, as when we directly predict the jointly target firms, we obtain roughly the same misclassification error as when we predict the constrained firms and the creditworthy firms separately. We therefore choose to keep the two predictions separate.

A4. Details on the forecasting of credit-constrained firms

The first exercise is designed to predict the credit-constrained firms by means of the ML algorithms described above.

In order to implement the first algorithm (decision tree), one needs to choose the complexity parameter c_p. We do this through cross-validation over the interval $[0.0005, \infty)$. As one can see from Figure A1, the optimal c_p is 0.00142.

Figure A2 shows the variables relative importance, a numeric value ranking the relative importance of variables. This includes not only variables that are primary splits and therefore are relevant for the final prediction (i.e. they appear in Figure A11 of Section A6), but also surrogate variables that, in some of the splits, would have done almost as well as the primary ones. In this way we also understand the role played by variables that are very highly correlated with those that appear in the final decision tree, although they do not actually show up as primary splits. The list and the order by which variables appear in the ranking do not necessarily correspond to that of the pruned tree graph of Figure A11. For instance, the variable ranked as first in Figure A2 may not be the variable chosen for the first split in Figure A11, and some variables not showing up in the pruned tree graph may be present in the relative importance graph. This happens because, given that a variable may appear many times in the tree, either as a primary or a surrogate splitting variable, its overall relative importance value is defined additively, as the

[21] After the under-sampling procedure the training set counts 75,777 observations (initially the training set contained 185,256 observations). The testing set remains the same.

sum of goodness of split measures for each split in which it was the primary variable, plus the sum of adjusted goodness measures for splits in which it was a surrogate.[22]

As one can see from Figure A2, the list of most important variables for the decision tree algorithm, in addition to those that already show up in the tree, includes: the total amount of loans granted to the firm at time t (variable = acco), the total amount of financial debts at $t-1$ (variable = debfin_Ly1), the dummy identifying firms that have at least one lending relationship with a total loan amount exceeding the €30,000 CR threshold (variable = no_aff), total assets at $t-1$ (variable = imm_Ly1) and the dummy identifying those firms that have already been a beneficiary of the GF guarantee in the past (variable = ben_FG_T0).

As for the second algorithm, the random forest, we need to choose the number of trees of the forest and the number of variables randomly selected for each tree. To do this, we look at the *out of bag error* (OOB) of the random forest. In Figure A3, we can see the OOB errors for the number of trees going from 1 to 500 and the number of variables going from 2 to 8. We choose the parameters with the lowest OOB error, that is: $n = 478$ and $m = 5$.

As expected, since the random forest is essentially the average of n not pruned decision trees, the list of important variables selected by the random forest algorithm contains the variables that are important in the decision tree. As we can see from Figure A4, in addition to all the variables already selected by the decision tree, other variables such as the age of the firm and Cerved rating of the firm in $t-1$ (rating_Ly1) also appear among the most important predictors.

As for the LASSO regression, we select the regularization parameter λ as to minimize the misclassification error according to the one-standard-error rule. Figure A5 shows the optimal λ chosen is 0.016 (whose logarithm is equal to -4.135), which is associated with the presence of six non-null coefficients in the regression model, shown in Table A5.[23]

A5. Details on forecasting creditworthy firms

As before, the complexity parameter c_p for the decision tree algorithm is chosen through cross-validation. The cross validation interval is $[0.001, \infty)$ as a result of the trade-off between the accuracy and interpretability of the resulting tree, in favor of a more readable tree structure. If we validate the c_p over a larger interval, we obtain a decision tree extremely hard to interpret and nevertheless dominated by other methods such as the random forest in terms of prediction accuracy. As one can see from Figure A6, the optimal c_p selected is 0.00141.

Figure A7 reports variables relative importance. As one can see, in addition to the splitting variables that already appeared in the tree graph (see Figure A12 in Section A6), the relative importance graph includes: a dummy variable (PNnull_Ly1) that describes a firm with

[22] The misclassification is used as a ranking criterion: each observation is classified using the best feasible surrogate rule.

[23] The sequence of λ parameters used in the cross-validation counts 600 values, generated by a sequence ranging within the interval $[0.2, 0.0005]$ with a uniform increment of 0.0005.

null equity or not and a dummy variable (PNneg_Ly1) that identifies firms with negative equity or not.

As for the random forest, Figure A8 shows the OOB errors graph, which allows us to choose the combination of number of trees and number of variables that minimizes such an error. As happened for the constrained firms forecasting, the important variables of the tree are important also for the random forest (Figure A9).

Before fitting the LASSO regression, we validate the penalizing parameter λ through 10-folds cross-validation. Figure A10 shows the cross validation error graph: the best λ selected according to the one-standard-error rule is equal to 0.00039 (whose logarithm is equal to -7.849), which is associated with the presence of 45 non-null coefficients in the estimated model, listed in Table A6.

A6. Model prediction results and model selection

An initial understanding of the characteristics of targeted firms can be provided by the decision tree, which is a good compromise between flexibility and interpretability. For this tool, the estimated (trained) algorithm essentially resembles a decision rule, in which each step (node) discriminates firms according to the value of a specific variable. Figure A11 shows the decision rule to predict credit-constrained firms, which tend to be those with few lending relationships with banks and a small variation in used credit, or those that have a larger number of lending relationships and greater exposure to total medium-long term debts. Figure A12 shows a more complicated prediction for creditworthy firms, which essentially depends on the Cerved-rating score, which is a balance-sheet summary of the firms' financial soundness, the presence of past defaults and exposure to the bank. In this case, also the past presence of a GF guarantee plays a role. The prediction from the random forest is less interpretable, as it combines many different trees. One can construct measures of variable importance, but we do not get a neat decision rule. The LASSO predictions are in principle more interpretable, but the presence of interactions and powers makes it less so. The difficulty in interpreting the forecasting rules raises some transparency concerns that we discuss in Subsection 6.2.

Our main aim is to have a forecasting rule that performs well out of sample. We therefore compare the models by looking at misclassification in the testing sample. The misclassification tables focus on the *false positive rate* (FP), which is the fraction of actually negative observations that are predicted as positive, and at the *false negative rate* (FN), which is the fraction of actually positive observations that are predicted as negative. Positive means that they are in the target status, and vice versa for negative. We define the predicted status as positive when the forecast probability of being so is larger than 0.5. For the credit-constrained exercise (Table A7), the decision tree and random forest performances are similar overall. The decision tree tends to do worse in classifying the actually non-constrained firms (as the FP rate is higher), while the random forest does worse in classifying the actually constrained. LASSO has a higher misclassification rate. For the creditworthy prediction (Table A8), the lowest misclassification

rate is reached by the random forest. In this case the decision tree has the worst performance and LASSO is in between.

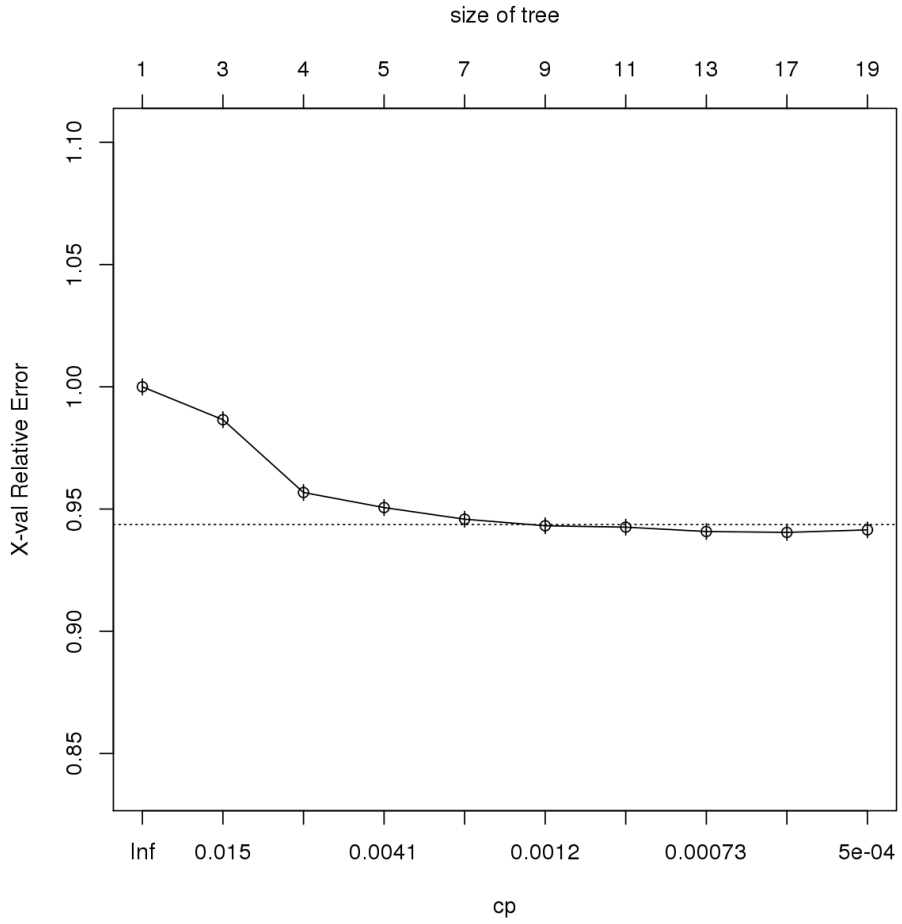
Comparing different models might be misleading if the total fraction of predicted positive cases is different across different algorithms.[24] Instead of classifying firms as target if the predicted probability is larger than 0.5, we can follow two approaches. The first approach orders all the observations according to the predicted probability and assigns to the target group the fraction x with the highest forecasted probability. In this way we can compare the algorithms performance keeping fixed the fraction of predicted positive cases. The Lift curve (Figure A13) looks at the how the *true positive rate* (TP, the fraction of actually positive observations that are forecasted as positive) changes with x (Hastie et al., 2009). For example, the point $x = 0.20$ means that the 20 per cent with the highest predicted probability of being in the target status is classified as 1 and all the rest as 0. The diagonal line is a random classifier (gives equal probability 0.5 to each observation): with this kind of classifier, at $x = 0.2$ one should predict correctly the 20 per cent of positive observations. If one uses a better classifier, she should expect to have more than 20 per cent of correctly predicted observations in the top 20 per cent of predicted probability. Again, the random forest does slightly better in both cases.

The second approach considers the entire set of possible thresholds that can be used to classify each observation as target or not. By changing the threshold, we obtain, for each algorithm, different combinations of the false positive rate (FP) and true positive rate (TP). The Receiver Operating Characteristic (ROC) curve shows all possible combinations for each algorithm (Hastie et al., 2009). Again, the diagonal line is a random classifier leading to equality between FP and TP rates. If one uses a better classifier, she should expect to have a TP rate higher than that obtained from the random classifier for each FP rate. This provides a graphical representation of the trade-off between the benefits of good positive classification and the costs implied by prediction errors. Looking at the ROC, the best classifier in both exercises is random forest (Figure A14).

[24] Furthermore, accuracy rates can be unreliable metrics of performance for unbalanced data sets: for example, if we imagine that we have an extremely unbalanced set with 95 per cent of red balls and 5 per cent of blue balls a totally red-classifier (predicting all balls red) will have high accuracy in terms of misclassification error (only 0.05) but it will nevertheless be completely useless.

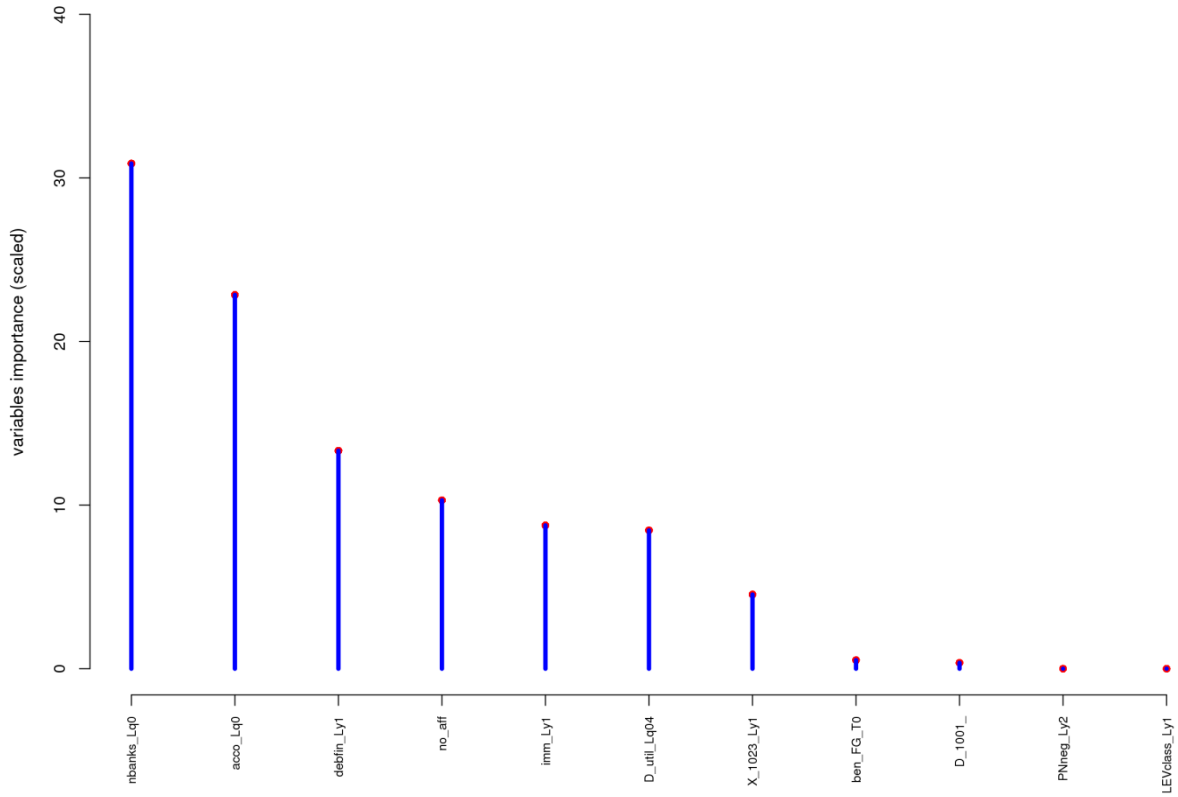
Appendix Figures

Figure A1. Complexity parameter validation of the tree for the credit-constrained exercise



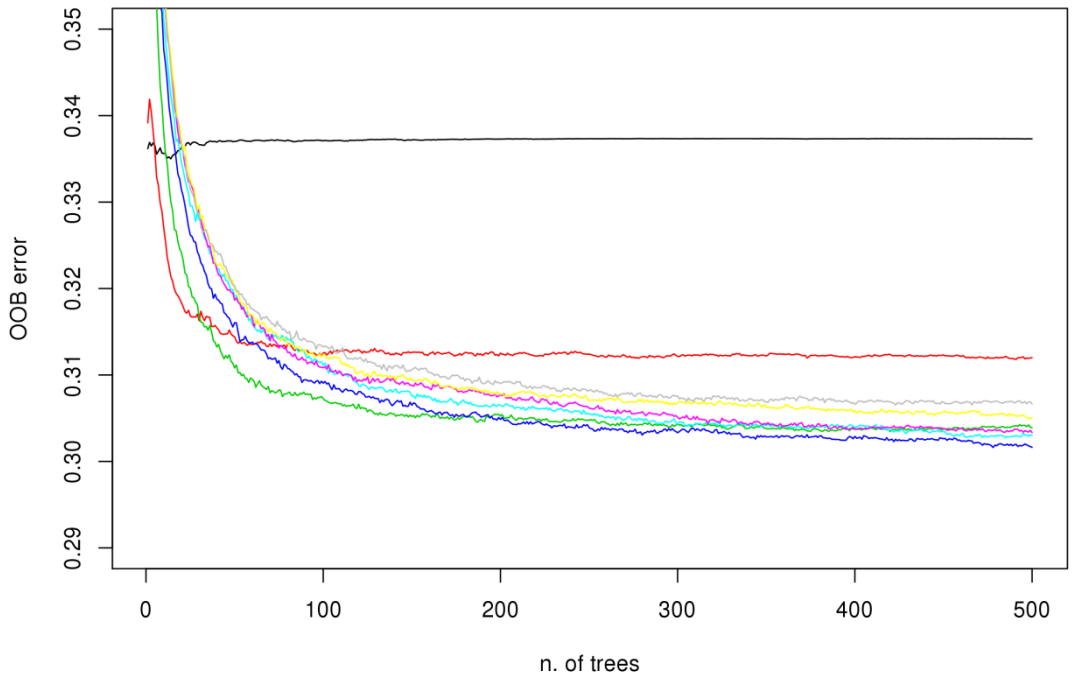
Notes. On the vertical axis the cross validation error of the tree is built with the correspondent complexity parameter on the horizontal axis.

Figure A2. Variables importance in the tree for the credit-constrained exercise



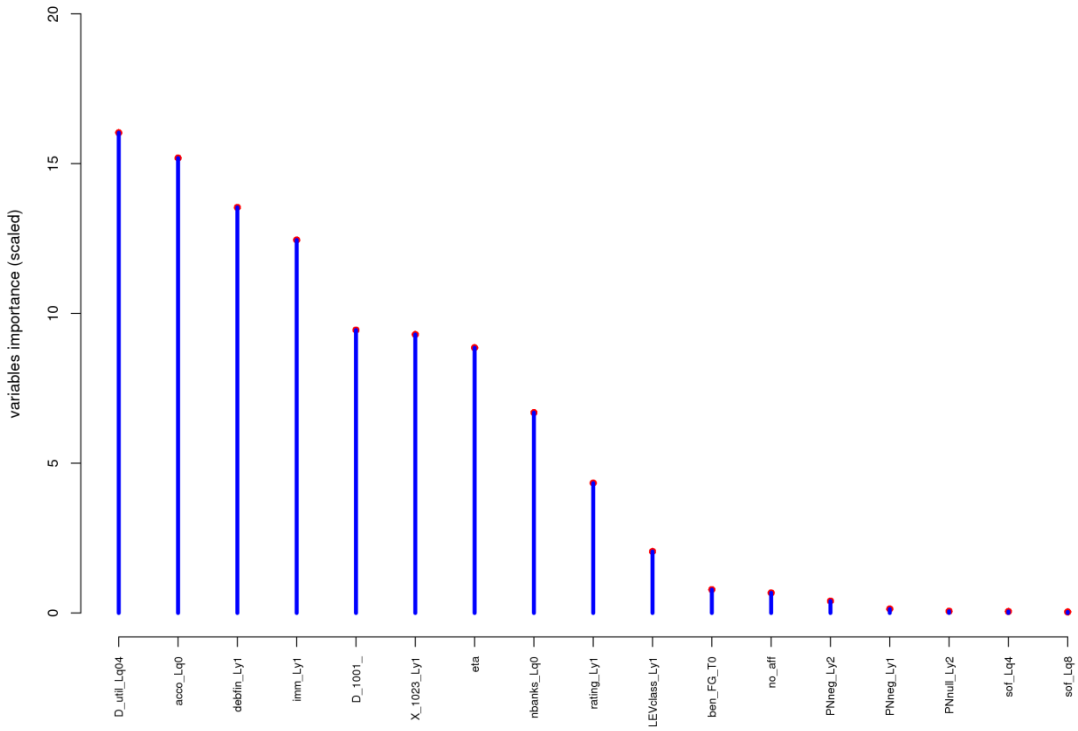
Notes. The vertical axis shows the scaled importance of variables in the tree. Variable description as follows. D_util_Lq04=Change in the total amount of bank loans granted and actually used by the firm, between the quarter when the PI request was issued and the same quarter in the previous year; acco_Lq0=Amount of total bank loans granted to the firm in the quarter in which the PI request is issued (Lq0); nbanks_Lq0=Number of banks lending money to the firm in the quarter in which the PI request is issued (Lq0); no_aff=Binary variable identifying whether data on firm's credit history is available (=1) or not (=0) in the CR dataset; D_1001_=Change in the variable X_1001_Ly1 (intangible assets) with respect to the previous year; X_1023_Ly1=Long term debts; ben_FG_T0=Binary variable identifying whether the firm has already been a beneficiary of the GF-guarantee program (=1) or not (=0) before the PI request was issued; imm_Ly1=Total assets (intangible + tangible assets) based on the most recent balance-sheet data available when the PI was issued (Ly1); PNneg_Ly2=Binary variable identifying whether the firm has negative equity (=1) or not (=0) lagged by 1 year with respect to PNneg_Ly1 (Ly2); debfin_Ly1=Total amount of short and long term debts, based on the most recent balance-sheet data available when the PI was issued; LEVclass_Ly1=Leverage class, based on the most recent balance-sheet data available when the PI was issued (Ly1).

Figure A3. Out of bag error of the random forest for the credit-constrained exercise



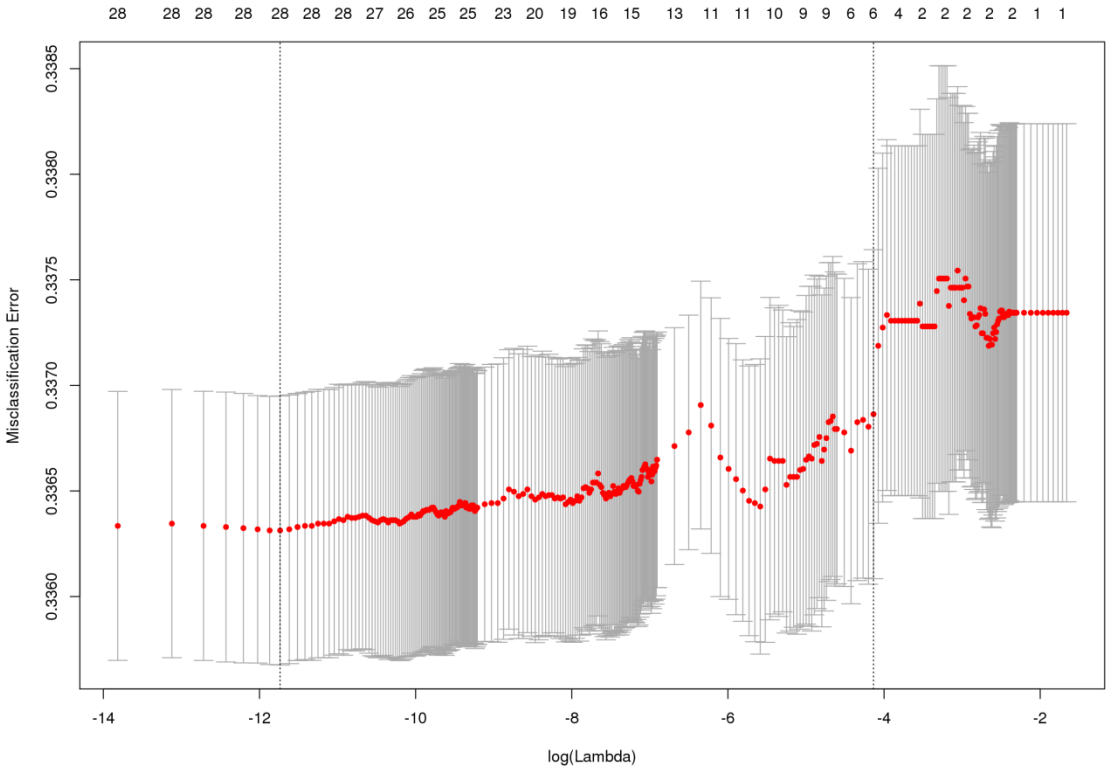
Notes. Each line in the graph corresponds to the random forest built with different numbers of variables. colors legend: black stands for 1 variable; red for 2 variables; green for 3 variables; blue for 4 variables; light blue for 5 variables; deep pink for 6 variables; yellow for 7 variables; light grey for 8 variables.

Figure A4. Variables importance in the random forest for the credit-constrained exercise



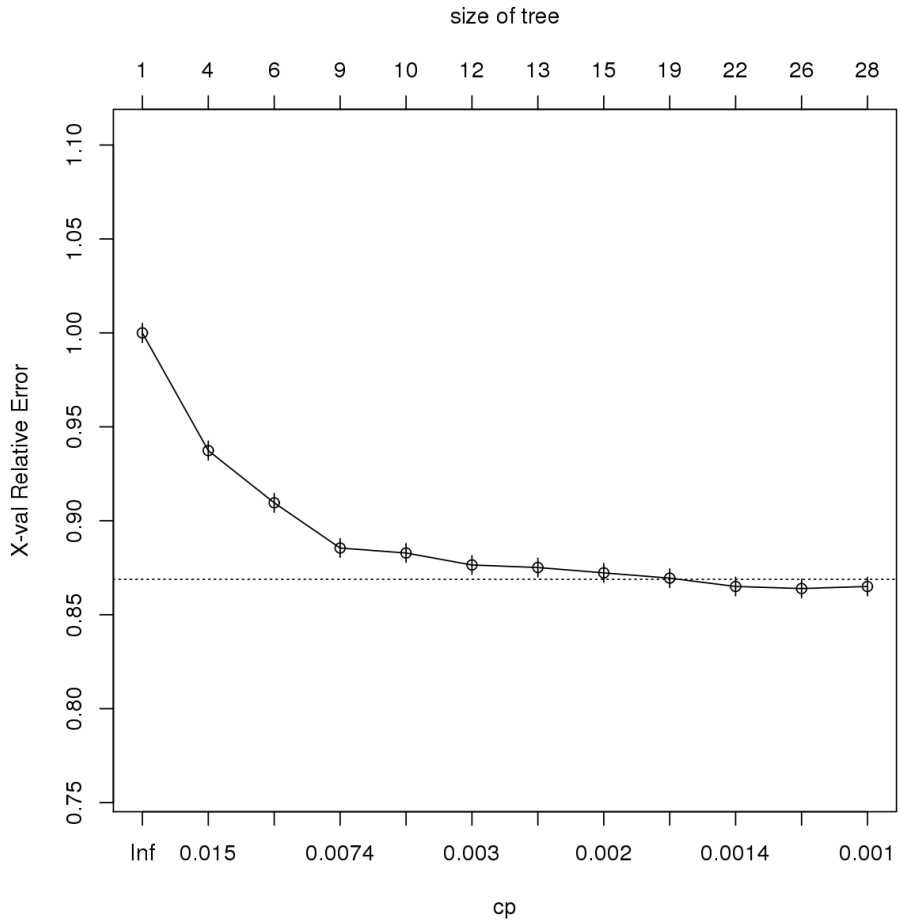
Notes. The vertical axis shows the scaled importance of variables in the random forest. Variable description as follows. D_util_Lq04=Change in the total amount of bank loans granted and actually used by the firm, between the quarter when the PI request was issued and the same quarter in the previous year.; acco_Lq0=Amount of total bank loans granted to the firm in the quarter in which the PI request is issued (Lq0); nbanks_Lq0=Number of banks lending money to the firm in the quarter in which the PI request is issued (Lq0); sof_Lq4=Binary variable identifying whether a firm has been reported to have bad loans (=1) or not (=0) 4 quarters before the PI request (Lq4); sof_Lq8=Binary variable defined as before but 8 quarters before (Lq8); no_aff=Binary variable identifying whether data on firm's credit history is available (=1) or not (=0) in the CR dataset; D_1001_=Change in the variable X_1001_Ly1 (intangible assets) with respect to the previous year; X_1023_Ly1=Long term debts; Eta=Firm age (expressed in years); ben_FG_T0=Binary variable identifying whether the firm has already been a beneficiary of the GF-guarantee program (=1) or not (=0) before the PI request was issued; rating_Ly1=Rating index produced by Cerved measuring firms' level of riskiness, based on the elaboration of balance-sheet data available when the PI is issued (Ly1); imm_Ly1=Total assets (intangible + tangible assets) based on the most recent balance-sheet data available when the PI was issued (Ly1); PNnull_Ly2=Binary variable identifying whether the firm has null equity (=1) or not (=0) lagged by 1 year with respect to PNnull_Ly1 (Ly2); PNneg_Ly1=Binary variable identifying whether the firm has negative equity (=1) or not (=0), based on the most recent balance-sheet data available when the PI was issued (Ly1); PNneg_Ly2=the same as before but lagged by 1 year with respect to PNneg_Ly1 (Ly2); debfin_Ly1=Total amount of short and long term debts, based on the most recent balance-sheet data available when the PI was issued; LEVclass_Ly1=Leverage class, based on the most recent balance-sheet data available when the PI was issued (Ly1).

Figure A5. Errors of the penalizing parameter for the credit-constrained exercise



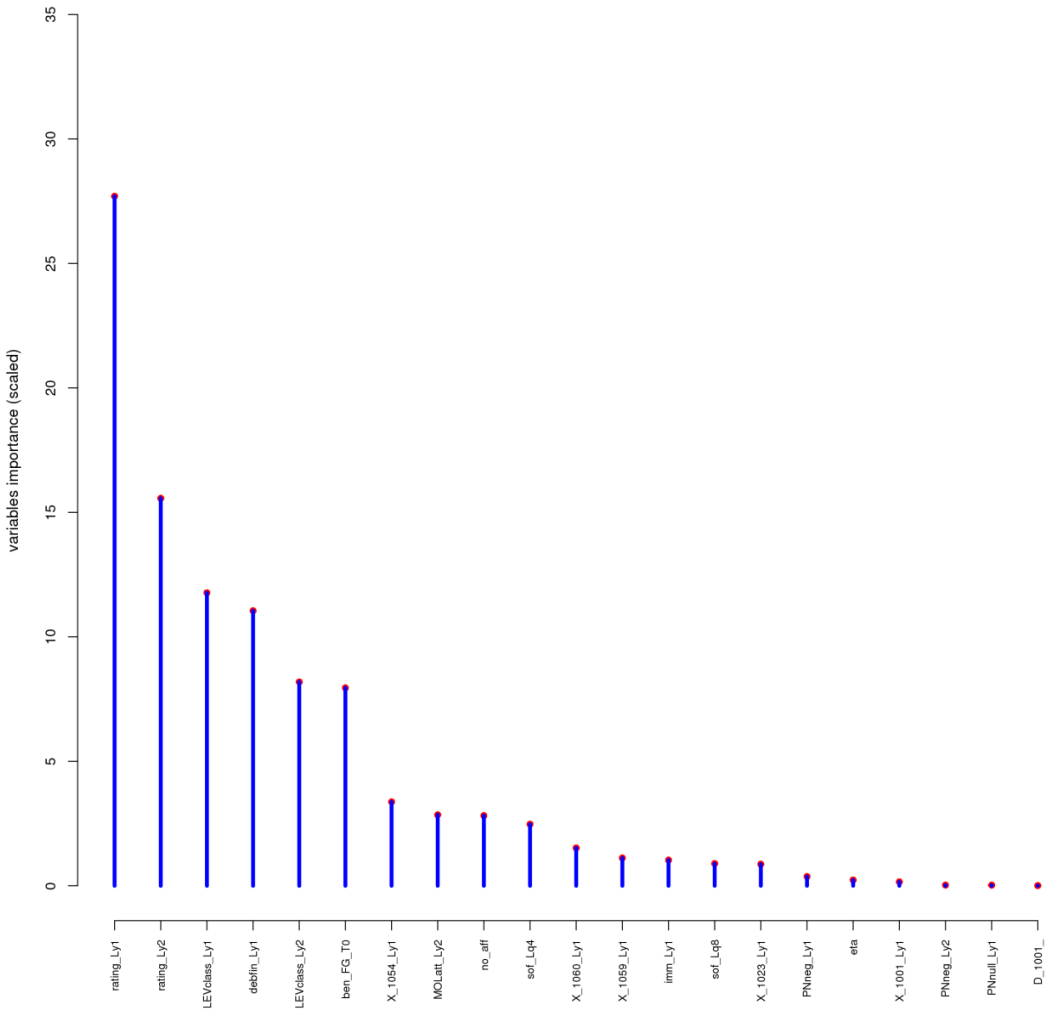
Notes. The graph shows the misclassification error (computed with cross validation) of regressions calculated using different penalizing parameters (on the bottom horizontal axis) and the number of nonzero coefficients (on the top horizontal axis).

Figure A6. Complexity parameter validation of the tree for the creditworthy exercise



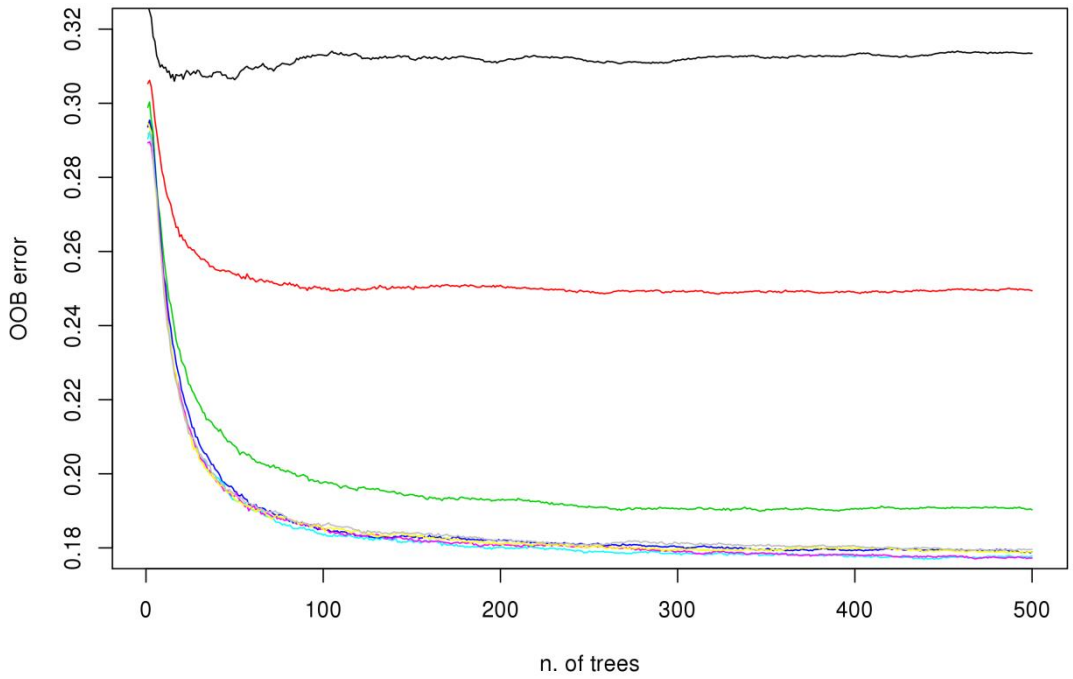
Notes. On the vertical axis the cross validation error of the tree is built with the correspondent complexity parameter on the horizontal axis.

Figure A7. Variables importance in the tree for the creditworthy exercise



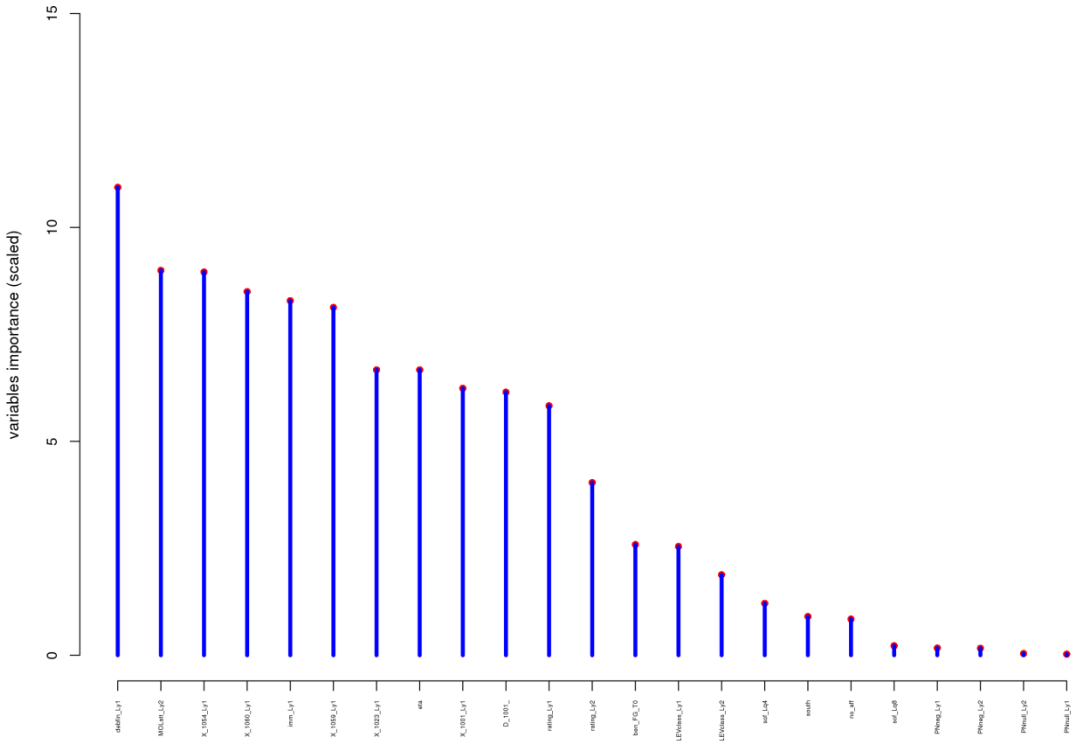
Notes. The vertical axis shows the scaled importance of variables in the tree. Variable description as follows. sof_Lq4=Binary variable identifying whether a firm has been reported to have bad loans (=1) or not (=0) in the CR 4 quarters before the PI request (Lq4); sof_Lq8=Binary variable defined as before but 8 quarters before the PI request (Lq8); no_aff=Binary variable identifying whether data on firm's credit history is available (=1) or not (=0) in the CR dataset, X_1001_Ly1=Intangible assets; D_1001_=Change in the variable X_1001_Ly1 with respect to the previous year; X_1023_Ly1=Long term debts; X_1054_Ly1=Production value; X_1059_Ly1=Labor cost; X_1060_Ly1=Gross operating margin; Eta=Firm age (expressed in years); ben_FG_T0=Binary variable identifying whether the firm has already been a beneficiary of the GF-guarantee program (=1) or not (=0) before the PI request was issued; rating_Ly1=Rating index produced by Cerved measuring firms' level of riskiness, based on the elaboration of balance-sheet data available when the PI is issued (Ly1); rating_ly2=the same as before but lagged by 1 year with respect to rating_Ly1 (Ly2); imm_Ly1=Total assets (intangible + tangible assets) based on the most recent balance-sheet data available when the PI was issued (Ly1); MOLatt_Ly2=Operating margin on assets index, based on the most recent balance-sheet data available when the PI was issued (Ly1) lagged by 1 year with respect to MOLatt_Ly1 (Ly2); PNnull_Ly1=Binary variable identifying whether the firm has null equity (=1) or not (=0) based on the most recent balance-sheet data available when the PI was issued (Ly1); PNneg_Ly1=Binary variable identifying whether the firm has negative equity (=1) or not (=0) based on the most recent balance-sheet data available when the PI was issued (Ly1); PNneg_Ly2=the same as before but lagged by 1 year with respect to PNneg_Ly1 (Ly2); debfin_Ly1=Total amount of short and long term debts, based on the most recent balance-sheet data available when the PI was issued; LEVclass_Ly1=Leverage class, based on the most recent balance-sheet data available when the PI was issued (Ly1); LEVclass_Ly2=same as before but lagged by one year (Ly2).

Figure A8. Out of bag error of the random forest for the creditworthy exercise



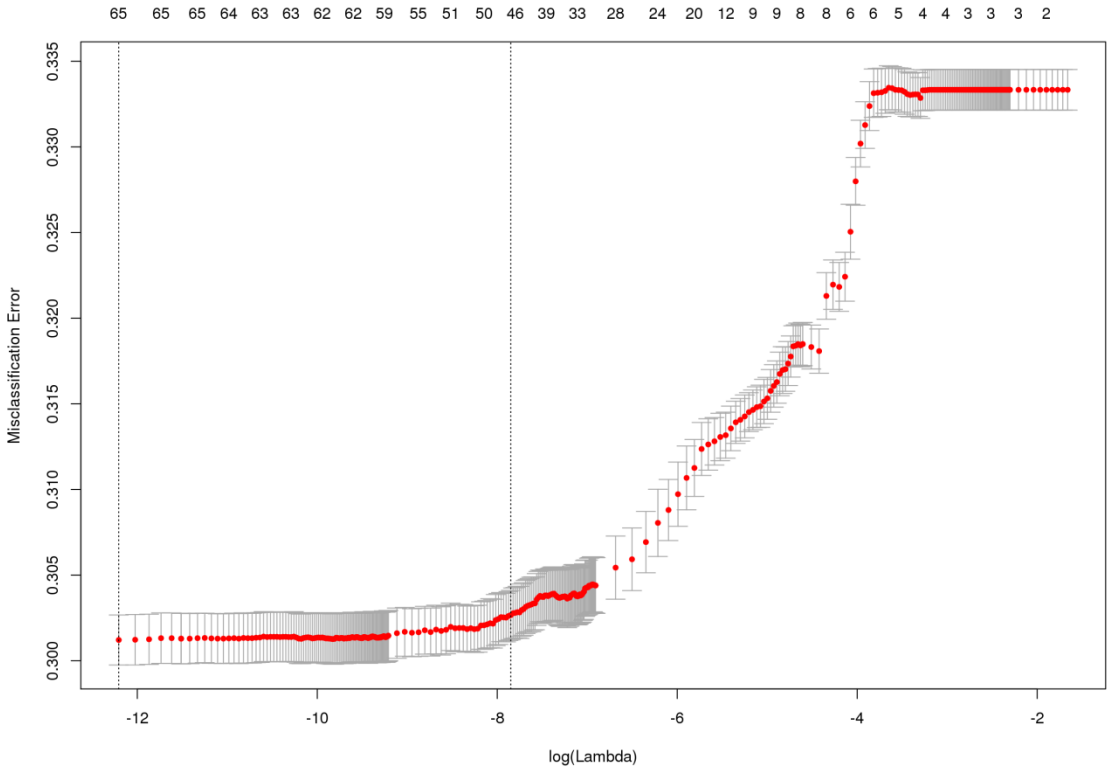
Notes. Each line in the graph corresponds to the random forest built with different numbers of variables. Colors legend: black stands for 1 variable; red for 2 variables; green for 3 variables; blue for 4 variables; light blue for 5 variables; deep pink for 6 variables; yellow for 7 variables; light grey for 8 variables.

Figure A9. Variables importance in the random forest for the creditworthy exercise



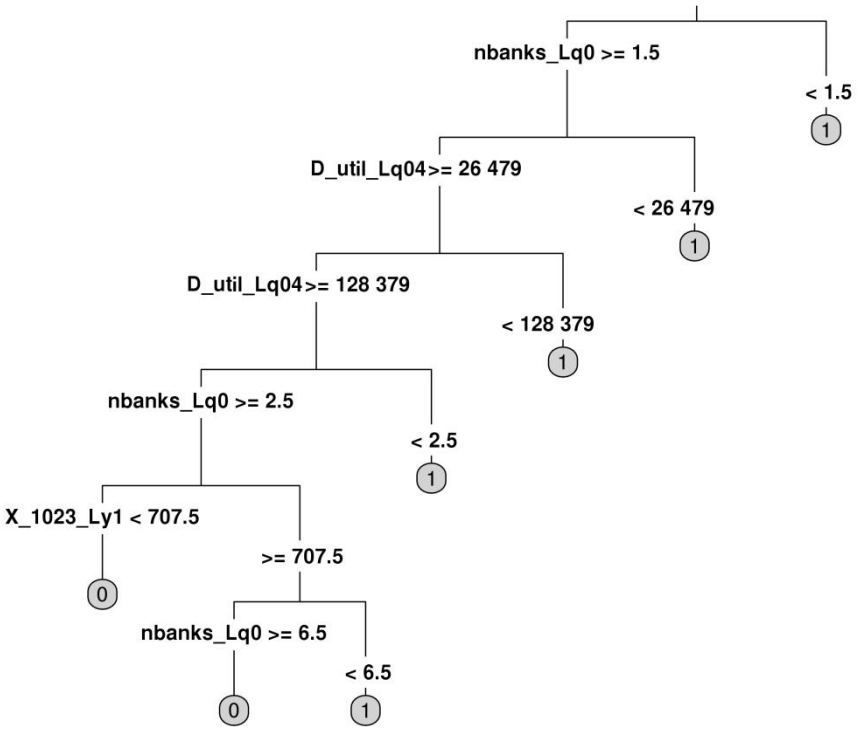
Notes. The vertical axis shows the scaled importance of variables in the random forest. Variable description as follows. sof_Lq4= Binary variable identifying whether a firm has been reported to have bad loans (=1) or not (=0) in the CR 4 quarters before the PI request (Lq4); sof_Lq8=Binary variable defined as before but 8 quarters before (Lq8); no_aff=Binary variable identifying whether data on firm's credit history is available (=1) or not (=0) in the CR dataset; X_1001_Ly1=Intangible assets; D_1001_=Change in the variable X_1001_Ly1 with respect to the previous year; X_1023_Ly1=Long term debts; X_1054_Ly1=Production value; X_1059_Ly1=Labor cost; X_1060_Ly1=Gross operating margin; Eta=Firm age (expressed in years); ben_FG_T0=Binary variable identifying whether the firm has already been a beneficiary of the GF-guarantee program (=1) or not (=0) before the PI request was issued; rating_Ly1=Rating index produced by Cerved measuring firms level of riskiness, based on the elaboration of balance-sheet data available when the PI is issued (Ly1); rating_Ly2=same as before but lagged by 1 year with respect to rating_Ly1 (Ly2); imm_Ly1=Total assets (intangible + tangible assets) based on the most recent balance-sheet data available when the PI was issued (Ly1); MOLatt_Ly2= operating margin on assets index based on the most recent balance-sheet data available when the PI was issued (Ly1) lagged by 1 year with respect to MOLatt_Ly1 (Ly2); PNnull_ly1=Binary variable identifying whether the firm has null equity (=1) or not (=0), based on the most recent balance-sheet data available when the PI was issued (Ly1); PNnull_ly2=same as before but lagged by 1 year with respect to PNnull_Ly1 (Ly2); PNneg_ly1=Binary variable identifying whether the firm has negative equity (=1) or not (=0), based on the most recent balance-sheet data available when the PI was issued (Ly1); PNneg_Ly2=same as before but lagged by 1 year with respect to PNneg_Ly1 (Ly2); debfin_Ly1=Total amount of short and long term debts, based on the most recent balance-sheet data available when the PI was issued; LEVclass_Ly1=Leverage class, based on the most recent balance-sheet data available when the PI was issued (Ly1); LEVclass_Ly2=same as before but lagged by 1 year with respect to LEVclass_ly1 (Ly2); South=Binary variable identifying whether the firm is located in the South of Italy (=1) or not (=0).

Figure A10. Errors of the penalizing parameter for the creditworthy exercise



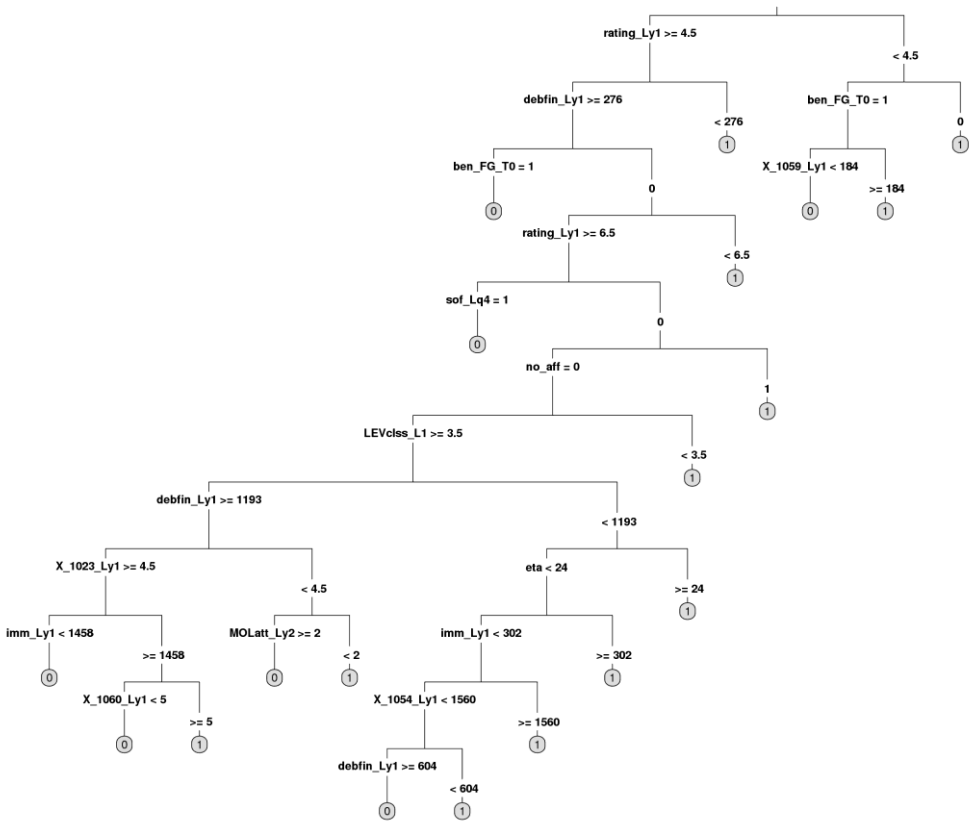
Notes. The graph shows the misclassification error (computed with cross validation) of regressions calculated using different penalizing parameters (on the bottom horizontal axis) and the number of nonzero coefficients (on the top horizontal axis).

Figure A11. Classification tree for the credit-constrained exercise



Notes. D_util_Lq04=Change in the total amount of bank loans granted and actually used by the firm, between the quarter when the PI request was issued and the same quarter in the previous year; nbanks_Lq0=Number of banks lending money to the firm in the quarter in which the PI request is issued (Lq0); X_1023_Ly1=Long term debts.

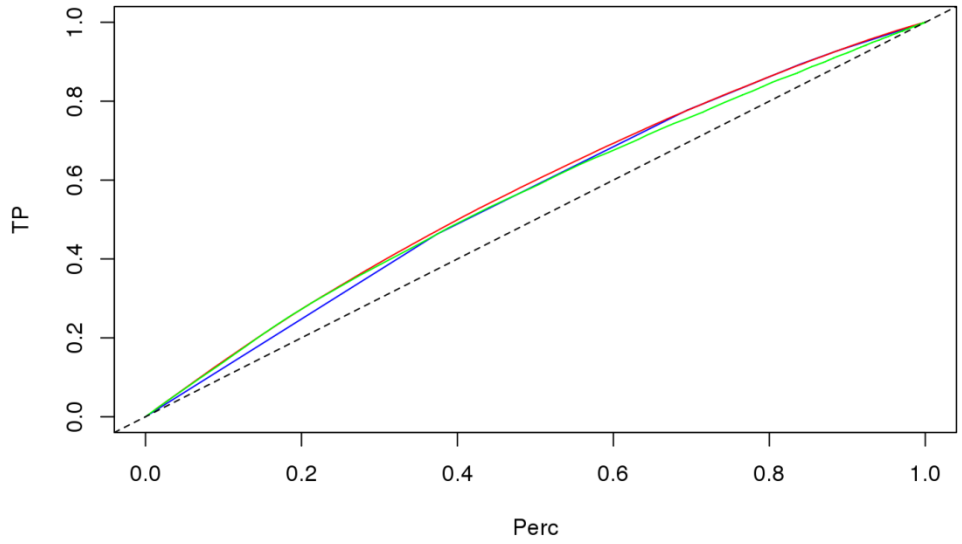
Figure A12. Classification tree for the creditworthy exercise



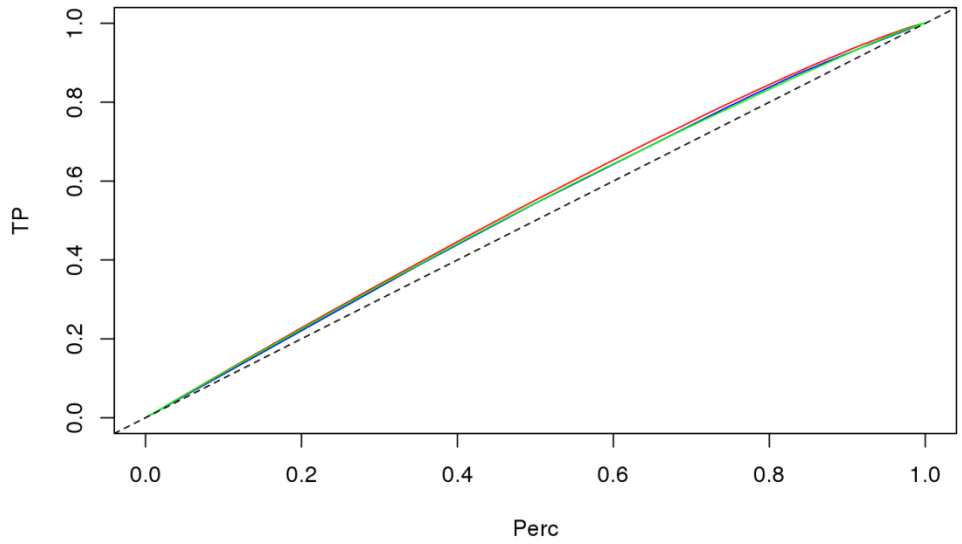
Notes. sof_Lq4=Binary variable identifying whether a firm has been reported to have bad loans (=1) or not (=0) to the CR 4 quarters before the PI request; no_aff=Binary variable identifying whether data on firm's credit history is available (=1) or not (=0) in the CR dataset; X_1023_Ly1=Long term debts; X_1054_Ly1=Production value; X_1059_Ly1=Labor cost; X_1060_Ly1=Gross operating margin (most recent balance-sheet data available when the PI was issued); eta=Firm age (expressed in years); ben_FG_T0=Binary variable identifying whether the firm has already been a beneficiary of the GF-guarantee program (=1) or not (=0) before the PI request was issued; rating_Ly1=Rating index produced by Cerved measuring firms' level of riskiness, based on the elaboration of balance-sheet data available when the PI is issued (Ly1); imm_Ly1=Total assets (intagible + tangible assets) based on the most recent balance-sheet data available when the PI was issued (Ly1); MOLatt_Ly2=Operating margin on assets index lagged by 1 year with respect to MOLatt_Ly1 (Ly2); debfin_Ly1=Total amount of short and long term debts based on the most recent balance-sheet data available when the PI was issued; LEVclass_Ly1=Leverage class, based on the most recent balance-sheet data available when the PI was issued (Ly1).

Figure A13. Lift curves

(a) Credit constrained



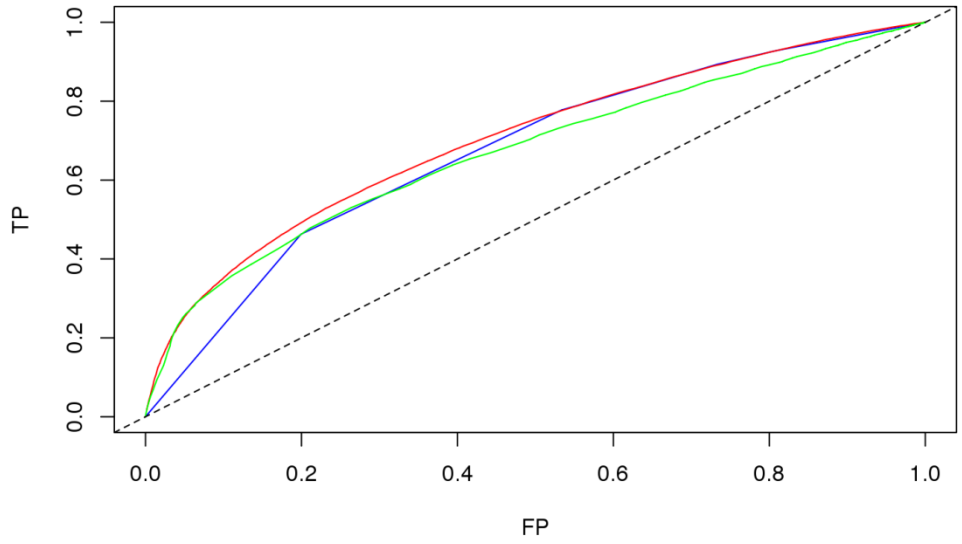
(b) Creditworthy



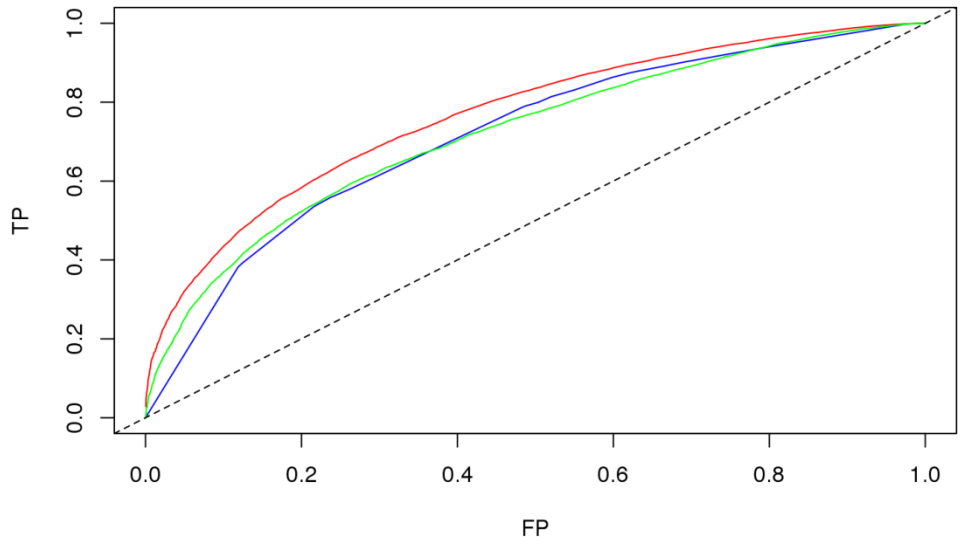
Notes. Testing sample (2011). The vertical axis shows the true positive ratio. On the horizontal axis the percentage of observations classified as positive, choosing first those with the highest predicted probability. Color legend: red is the random forest Lift curve; blue is the decision tree Lift curve; green is the LASSO Lift curve.

Figure A14. ROC curves

(a) Credit constrained



(b) Creditworthy



Notes. Testing sample (2011). The vertical axis shows the true positive ratio. The horizontal axis shows the false positive ratio. Color legend: red is the random forest ROC curve; blue is the decision tree ROC curve; green is the LASSO ROC curve.

Appendix Tables

Table A1. Complete list of variables and brief description

Variable	Source	Description
draz	CR (elaboration)	Binary response variable identifying whether a firm is constrained (=1) or not (=0)
credit_worthy	CR (elaboration)	Binary response variable identifying whether a firm is creditworthy (=1) or not (=0)
util_Lq0	CR	Amount of bank loans granted and actually used by the firm in the quarter in which the PI request is issued (Lq0)
util_Lq1	CR	Variable util_Lq0 lagged by 1 quarter (Lq1)
util_Lq2	CR	Variable util_Lq0 lagged by 2 quarters (Lq2)
util_Lq3	CR	Variable util_Lq0 lagged by 3 quarters (Lq3)
util_Lq5	CR	Variable util_Lq0 lagged by 5 quarters (Lq5)
util_Lq6	CR	Variable util_Lq0 lagged by 6 quarters (Lq6)
util_Lq7	CR	Variable util_Lq0 lagged by 7 quarters (Lq7)
D_util_Lq04	CR	Change in the total amount of bank loans granted and actually used by the firm, between the quarter when the PI request was issued and the same quarter in the previous year
D_util_Lq08	CR	Change in the total amount of bank loans granted and actually used by the firm, between the quarter when the PI request was issued and the same quarter two years earlier
acco_Lq0	CR	Amount of total bank loans granted to the firm in the quarter in which the PI request is issued (Lq0)
acco_Lq1	CR	Variable acco_Lq0 lagged by 1 quarter (Lq1)
acco_Lq2	CR	Variable acco_Lq0 lagged by 2 quarters (Lq2)
acco_Lq3	CR	Variable acco_Lq0 lagged by 3 quarters (Lq3)
acco_Lq5	CR	Variable acco_Lq0 lagged by 5 quarters (Lq5)
acco_Lq6	CR	Variable acco_Lq0 lagged by 6 quarters (Lq6)
acco_Lq7	CR	Variable acco_Lq0 lagged by 7 quarters (Lq7)
D_acco_Lq04	CR	Change in the total amount of loans granted to the firm, between the quarter when the PI request was issued and the same quarter in the previous year
D_acco_Lq08	CR	Change in the total amount of loans granted to the firm, between the quarter when the PI request was issued and the same quarter two years earlier
nbanks_Lq0	CR	Number of banks lending money to the firm in the quarter in which the PI request is issued (Lq0)
nbanks_Lq1	CR	Variable nbanks_Lq0 lagged by 1 quarter (Lq1)
		(continued)

		(continued)
nbanks_Lq2	CR	Variable nbanks_Lq0 lagged by 2 quarters (Lq2)
nbanks_Lq3	CR	Variable nbanks_Lq0 lagged by 3 quarters (Lq3)
D_nbanksLq04	CR	Change in the total number of banks lending money to the firm, between the quarter when the PI request was issued and the same quarter in the previous year
sof_Lq0	CR	Binary variable identifying whether a firm has been reported to have bad loans (=1) or not (=0) in the CR in the quarter in which the PI request is issued (Lq0). A firm has bad loans if she is reported as insolvent by any bank, regardless of the amount of loans borrowed from that bank
sof_Lq1	CR	Variable sof_Lq0 lagged by 1 quarter (Lq1)
sof_Lq2	CR	Variable sof_Lq0 lagged by 2 quarters (Lq2)
sof_Lq3	CR	Variable sof_Lq0 lagged by 3 quarters (Lq3)
sof_Lq4	CR	Variable sof_Lq0 lagged by 4 quarters (Lq4)
sof_Lq5	CR	Variable sof_Lq0 lagged by 5 quarters (Lq4)
sof_Lq6	CR	Variable sof_Lq0 lagged by 5 quarters (Lq5)
sof_Lq7	CR	Variable sof_Lq0 lagged by 6 quarters (Lq6)
sof_Lq8	CR	Variable sof_Lq0 lagged by 7 quarters (Lq7)
no_aff	CR	Binary variable identifying whether data on firm's credit history is available (=1) or not (=0) in the CR dataset
X_1001_Ly1	Cerved	Intangible assets
D_1001_	Cerved	Change in the variable X_1001_Ly1 with respect to the previous year
X_1002_Ly1	Cerved	Tangible fixed assets
D_1002_	Cerved	Change in the variable X_1002_Ly1 with respect to the previous year
X_1005_Ly1	Cerved	Total fixed asset
D_1005_	Cerved	Change in the variable X_1005_Ly1 with respect to the previous year
X_1014_Ly1	Cerved	Total short-term assets
D_1014_	Cerved	Change in the variable X_1014_Ly1 with respect to the previous year
X_1015_Ly1	Cerved	(Total) assets
D_1015_	Cerved	Change in the variable X_1015_Ly1 with respect to the previous year
X_1016_Ly1	Cerved	Shareholders' capital
D_1016_	Cerved	Change in the variable X_1016_Ly1 with respect to the previous year
X_1020_Ly1	Cerved	Equity
D_1020_	Cerved	Change in the variable X_1020_Ly1 with respect to the previous year
		(continued)

		(continued)
X_1021_Ly1	Cerved	Provisions
D_1021_	Cerved	Change in the variable X_1021_Ly1 with respect to the previous year
X_1023_Ly1	Cerved	Long term debts
D_1023_	Cerved	Change in the variable X_1023_Ly1 with respect to the previous year
X_1024_Ly1	Cerved	Long term debts towards banks
D_1024_	Cerved	Change in the variable X_1024_Ly1 with respect to the previous year
X_1047_Ly1	Cerved	Long term debts: other financial liabilities
D_1047_	Cerved	Change in the variable X_1047_Ly1 with respect to the previous year
X_1027_Ly1	Cerved	Short term debts towards banks
D_1027_	Cerved	Change in the variable X_1027_Ly1 with respect to the previous year
X_1048_Ly1	Cerved	Short term debts: other financial liabilities
D_1048_	Cerved	Change in the variable X_1048_Ly1 with respect to the previous year
X_1033_Ly1	Cerved	Short-term total liabilities
D_1033_	Cerved	Change in the variable X_1033_Ly1 with respect to the previous year
X_1034_Ly1	Cerved	Liabilities, net of advances received
D_1034_	Cerved	Change in the variable X_1034_Ly1 with respect to the previous year
X_1051_Ly1	Cerved	Net revenues
D_1051_	Cerved	Change in the variable X_1051_Ly1 with respect to the previous year
X_1054_Ly1	Cerved	Production value
D_1054_	Cerved	Change in the variable X_1054_Ly1 with respect to the previous year
X_1058_Ly1	Cerved	Operating value added
D_1058_	Cerved	Change in the variable X_1058_Ly1 with respect to the previous year
X_1059_Ly1	Cerved	Labor cost
D_1059_	Cerved	Change in the variable X_1059_Ly1 with respect to the previous year
X_1060_Ly1	Cerved	Gross operating margin
D_1060_	Cerved	Change in the variable X_1060_Ly1 with respect to the previous year
X_1067_Ly1	Cerved	Net financial income
D_1067_	Cerved	Change in the variable X_1067_Ly1 with respect to the previous year
X_1068_Ly1	Cerved	Current profit before financial charges in the current year
D_1068_	Cerved	Change in the variable X_1068_Ly1 with respect to the previous year
		(continued)

		(continued)
X_1069_Ly1	Cerved	Financial charges
D_1069_	Cerved	Change in the variable X_1069_Ly1 with respect to the previous year
X_1073_Ly1	Cerved	Taxes
D_1073_	Cerved	Change in the variable X_1073_Ly1 with respect to the previous year
X_1074_Ly1	Cerved	Net adjusted income
D_1074_	Cerved	Change in the variable X_1074_Ly1 with respect to the previous year
X_1076_Ly1	Cerved	Profit (Loss)
D_1076_	Cerved	Change in the variable X_1076_Ly1 with respect to the previous year
X_1026_Ly1	Cerved	Short term debts
D_1026_	Cerved	Change in the variable X_1026_Ly1 with respect to the previous year
eta	Cerved	Firm age (expressed in years)
ben_FG_T0	GF dataset	Binary variable identifying whether the firm has already been a beneficiary of the GF-guarantee program (=1) or not (=0) before the PI request was issued
rating_Ly1	Cerved	Rating index produced by Cerved measuring firms' level of riskiness, based on the elaboration of balance-sheet data: the index ranges from 1 to 9, higher values are associated to higher risk. The index refers to the most recent balance-sheet data available when the PI is issued (Ly1)
rating_Ly2	Cerved	Variable rating_Ly1 lagged by 1 year (Ly2)
imm_Ly1	Cerved (elaboration)	Total assets (intangible + tangible assets); it is based on the most recent balance-sheet data available when the PI was issued (Ly1)
imm_Ly2	Cerved (elaboration)	Variable imm_Ly1 lagged by 1 year (Ly2)
roa_Ly1	Cerved (elaboration)	Return on assets index; it is based on the most recent balance-sheet data available when the PI was issued (Ly1)
roa_Ly2	Cerved (elaboration)	Variable roa_Ly1 lagged by 1 year (Ly2)
MOLatt_Ly1	Cerved (elaboration)	Operating margin on assets index; it is based on the most recent balance-sheet data available when the PI was issued (Ly1)
MOLatt_Ly2	Cerved (elaboration)	Variable MOLatt_Ly1 lagged by 1 year (Ly2)
PNnull_Ly1	Cerved (elaboration)	Binary variable identifying whether the firm has null equity (=1) or not (=0); it is based on the most recent balance-sheet data available when the PI was issued (Ly1)
PNnull_Ly2	Cerved (elaboration)	Variable PNnull_Ly1 lagged by 1 year (Ly2)
		(continued)

		(continued)
PNneg_Ly1	Cerved (elaboration)	Binary variable identifying whether the firm has negative equity (=1) or not (=0); it is based on the most recent balance-sheet data available when the PI was issued (Ly1)
PNneg_Ly2	Cerved (elaboration)	Variable PNneg_ly1 lagged by 1 year (Ly2)
debfin_Ly1	Cerved (elaboration)	Total amount of short and long term debts; it is based on the most recent balance-sheet data available when the PI was issued (debfin_Ly1= X_1024_Ly1 + X_1027_Ly1 + X_1047_Ly1 + X_1048_Ly1)
debfin_Ly2	Cerved (elaboration)	Variable debfin_Ly1 lagged by 1 year (Ly2)
defret_t	Cerved (elaboration)	Binary variable identifying whether the firm has been reported to be in an adjusted default status (=1) or not (=0) to the CR, in the year in which the PI request is issued.
LEVclass_Ly1	Cerved (elaboration)	Leverage class, based on the most recent balance-sheet data available when the PI was issued (ly1). Leverage classes are defined as: Class 1: $25<LEV\_Ly1\leq 50$; Class 2: $0\leq LEV\_Ly1\leq 25$; Class 3: $50<LEV\_Ly1\leq 75$; Class 4: $75<LEV\_Ly1\leq 100$; Class 5: $LEV\_Ly1<0$ or $LEV\_Ly1>100$. The Leverage index is obtained as the ratio between total debts and the sum of total debts and equity, i.e. LEV_Ly1=debfin_Ly1/(debfin_Ly1+_1020_Ly1)
LEVclass_Ly2	Cerved (elaboration)	Variable LEVclass_Ly1 lagged by 1 year (Ly2)
South	Cerved (elaboration)	Binary variable identifying whether the firm is located in the South of Italy (=1) or not (=0)
Industria	Cerved (elaboration)	Binary variable identifying whether the firm works in the industrial cluster (=1) or not (=0) according to the ATECO07 classification rules

Table A2. Summary statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
draz	278,355	0.662291	0.4729296	0	1
credit_worthy	278,355	0.8637352	0.3430702	0	1
nbanks_Lq0	278,355	3.500742	3.925006	0	65
nbanks_Lq1	278,355	3.381549	3.886809	0	63
nbanks_Lq2	278,355	3.379946	3.868736	0	65
nbanks_Lq3	278,355	3.387832	3.862846	0	68
rating_Ly1	278,355	5.139926	1.950705	1	9
rating_Ly2	278,355	5.16766	1.946496	1	9
X_1001_Ly1	278,355	759.3384	33890.19	0	8006477
X_1016_Ly1	278,355	1382.7	26958.24	0	8515841
X_1020_Ly1	278,355	3057.165	37774.63	-805292	7137686
X_1058_Ly1	278,355	1972.476	30238.87	-248055	7929319
X_1059_Ly1	278,355	1303.892	19034.31	0	5689109
X_1069_Ly1	278,355	149.6218	3088.111	0	764986
X_1076_Ly1	278,355	95.99631	7723.9	-1243793	1765924
X_1021_Ly1	278,355	327.778	9691.388	0	2481209
X_1067_Ly1	278,355	81.77763	3248.203	-616209	1180472
X_1073_Ly1	278,355	150.8776	3197.765	-161024	944818
X_1068_Ly1	278,355	410.9461	10553.39	-1201972	2854104
X_1074_Ly1	278,355	95.7519	7717.045	-1243793	1759069
X_1060_Ly1	278,355	668.5834	14991.26	-440237	3862118
X_1024_Ly1	278,355	1039.929	18104.29	0	4145568
X_1047_Ly1	278,355	57.51599	2282.624	0	759100
X_1048_Ly1	278,355	19.51724	25.58316	0	3694.11
sof_Lq0	278,355	0.0152791	0.1226607	0	1
sof_Lq1	278,355	0.011956	0.1086879	0	1
sof_Lq2	278,355	0.0101489	0.1002295	0	1
sof_Lq3	278,355	0.0087335	0.0930441	0	1
sof_Lq4	278,355	0.007609	0.0868972	0	1
sof_Lq5	278,355	0.0065492	0.0806618	0	1
sof_Lq6	278,355	0.0059025	0.076601	0	1
sof_Lq7	278,355	0.0052271	0.0721099	0	1
sof_Lq8	278,355	0.0046919	0.0683363	0	1
eta	278,355	15.72364	12.54063	1	158
no_aff	278,355	0.130427	0.3367732	0	1
ben_FG_T0	278,355	0.1046901	0.3061542	0	1
roa_Ly1	278,355	2.168137	166.7608	-78600	2644
roa_Ly2	278,355	3.911367	32.66447	-2500	13400
MOLatt_Ly1	278,355	6.63502	132.5956	-59000	2650
MOLatt_Ly2	278,355	8.254592	34.88289	-2480	13400
PNnull_Ly1	278,355	0.0022741	0.0476331	0	1
PNnull_Ly2	278,355	0.0027339	0.0522155	0	1
PNneg_Ly1	278,355	0.0610156	0.2393594	0	1
PNneg_Ly2	278,355	0.0575811	0.2329501	0	1
defret_t	278,355	0.0573548	0.2325198	0	1
LEVclass_Ly1	278,355	3.434046	1.043287	1	5
LEVclass_Ly2	278,355	3.437941	1.034682	1	5
South	278,355	0.2081766	0.4060046	0	1
Industria	278,355	0.2875608	0.4526261	0	1
D_nbanksLq04	278,355	0.1170053	1.21296	-29	24
D_1001_	278,355	25.71761	7763.913	-664320	3417254
D_1002_	278,355	37.75182	8490.817	-3238995	846731
D_1005_	278,355	123.612	10507.52	-2390097	1117130
D_1014_	278,355	184.3684	12314.42	-3044824	1385245
					(continued)

					(continued)
D_1015_	278,355	307.9804	15884.79	-3051361	1836069
D_1016_	278,355	54.58276	6189.853	-2264168	1050000
D_1020_	278,355	105.1717	11770.97	-2370894	3875926
D_1026_	278,355	206.9517	15666.25	-2262240	4666332
D_1033_	278,355	101.0287	14764.14	-3388997	1403034
D_1034_	278,355	307.9804	15884.79	-3051361	1836069
D_1051_	278,355	-101.6028	38174.74	-9042206	8919936
D_1054_	278,355	-111.2534	38372.79	-9042206	8919936
D_1058_	278,355	19.69632	4954.837	-799439	542711
D_1059_	278,355	34.52865	1935.915	-194598	504550
D_1069_	278,355	-41.947	1238.145	-367852	147223
D_1076_	278,355	-15.43817	7101.941	-1170543	1619477
D_1021_	278,355	18.61368	2263.007	-389231	828918
D_1067_	278,355	-23.37891	3321.527	-1029975	304159
D_1073_	278,355	-2.192438	1625.085	-252359	500797
D_1074_	278,355	-14.7069	7093.923	-1170543	1619477
D_1060_	278,355	-14.83233	4283.792	-701717	610715
D_1024_	278,355	33.70418	7625.363	-1701778	1173540
D_1047_	278,355	6.703725	3721.354	-903000	757749.8
D_1027_	278,355	94.8289	14571.34	-3383334	1403034
D_1048_	278,355	.6057571	19.13432	-1062.5	3133.4
D_1023_	278,355	88.75335	10911.99	-1737727	2680583
D_1068_	278,355	-60.80531	5795.715	-1323831	364816
util_Lq0	278,355	2903870.75	20827374	0	2654373632
util_Lq1	278,355	2873342.25	20780410	0	3019080448
util_Lq2	278,355	2835801.25	21281142	0	3801057536
util_Lq3	278,355	2796632.25	21530432	0	3942093056
util_Lq5	278,355	2734884.75	22740976	0	4270024192
util_Lq6	278,355	2720034.5	23230754	0	4263706368
util_Lq7	278,355	2704553.5	23754744	0	4270024192
acco_Lq0	278,355	4490085.5	31087684	0	3998619648
acco_Lq1	278,355	4467964.5	30979026	0	4017632512
acco_Lq3	278,355	4462579.5	31500472	0	4546283520
acco_Lq2	278,355	4462436	31043964	0	4522943488
acco_Lq5	278,355	4468116	33159358	0	4900827648
acco_Lq6	278,355	4475285.5	33659944	0	4897276416
acco_Lq7	278,355	4480276	34099640	0	4945260032
X_1002_Ly1	278,355	2672.047363	48989.71094	0	14006136
X_1005_Ly1	278,355	4727.964355	83007.60156	0	15737555
X_1014_Ly1	278,355	6312.870605	130269.8359	0	45944344
X_1015_Ly1	278,355	11040.83496	178823.25	1	52464192
X_1026_Ly1	278,355	5394.856934	79528.22656	-284540	14704909
X_1027_Ly1	278,355	5491.078613	125194.7109	0	43820324
X_1033_Ly1	278,355	5645.978027	127866.9453	0	44739616
X_1034_Ly1	278,355	11040.83496	178823.25	1	52464192
X_1051_Ly1	278,355	10196.03418	136811.5938	0	31494924
X_1054_Ly1	278,355	10254.43848	137088.9688	-9805	31494924
X_1023_Ly1	278,355	1728.243896	43823.20313	0	10518466
imm_Ly1	278,355	3431.385742	69347.39844	0	14187642
imm_Ly2	278,355	3367.91626	68829.82813	0	14085695
debfin_Ly1	278,355	6608.040527	129915.3359	0	44070324
debfin_Ly2	278,355	6472.198242	130350.25	0	43878748
D_util_Lq04	278,355	147899.1406	7653159	-1778668928	977245184
D_util_Lq08	278,355	197887.8594	12428891	-2410919680	1261780608
D_acco_Lq04	278,355	22902.0957	9866105	-2004762112	1467343744
D_acco_Lq08	278,355	-4529.17041	14285915	-2588611328	1484282240

Table A3. List of variables after the screening for the credit-constrained exercise

Variable name
acco_Lq0
ben_FG_T0
D_1001_
debfin_Ly1
D_util_Lq04
eta
imm_Ly1
LEVclass_Ly1
nbanks_Lq0
no_aff
PNneg_Ly1
PNneg_Ly2
PNnull_Ly2
rating_Ly1
sof_Lq4
sof_Lq8
X_1023_Ly1

Table A4. List of variables after the screening for the creditworthy exercise

Variable
rating_Ly1
rating_Ly2
X_1001_Ly1
X_1054_Ly1
X_1059_Ly1
X_1060_Ly1
X_1023_Ly1
sof_Lq4
sof_Lq8
eta
no_aff
ben_FG_T0
imm_Ly1
MOLatt_Ly2
PNnull_Ly1
PNnull_Ly2
PNneg_Ly1
PNneg_Ly2
debfin_Ly1
LEVclass_Ly1
LEVclass_Ly2
South
D_1001_

Table A5. Coefficients of LASSO regression for the credit-constrained exercise

Variable	Coef.
imm_Ly1×debfin_Ly1	-0.000038241057647
acco_Lq0×debfin_Ly1	-0.000034496744523
nbanks_Lq0×imm_Ly1	0.001697873471909
rating_Ly1	0.024015592002447
nbanks_Lq0	-0.331841806204264
no_aff	0.436245372314423

Table A6. Coefficients of LASSO regression for the creditworthy exercise

Variable	Coef.
sof_Lq8	-1.491700
sof_Lq4×PNneg_Ly1	-1.220300
ben_FG_T0	-1.083400
sof_Lq4×south	-0.864931
rating_Ly2	-0.625037
PNneg_Ly1	-0.423501
rating_Ly1×south	-0.308869
LEVclass_Ly2	-0.241184
rating_Ly1×ben_FG_T0	-0.225360
rating_Ly2×eta	-0.204530
South	-0.124396
sof_Lq4×LEVclass_Ly2	-0.111410
rating_Ly1×LEVclass_Ly2	-0.066829
no_aff×LEVclass_Ly1	-0.065553
rating_Ly1×eta	-0.062732
rating_Ly2×LEVclass_Ly2	-0.052011
rating_Ly1×debfin_Ly1	-0.051245
rating_Ly2×debfin_Ly1	-0.045650
PNneg_Ly1×debfin_Ly1	-0.043521
no_aff×LEVclass_Ly2	-0.041634
debfin_Ly1	-0.021730
eta×south	-0.017883
rating_Ly2×PNneg_Ly2	0.001894
rating_Ly2×imm_Ly1	0.009092
imm_Ly1	0.011241
X_1059_Ly1×imm_Ly1	0.011459
X_1060_Ly1×debfin_Ly1	0.011492
PNneg_Ly1×south	0.018185
eta×ben_FG_T0	0.022311
rating_Ly1×no_aff	0.040397
rating_Ly2×sof_Lq8	0.043982
imm_Ly1×MOLatt_Ly2	0.063558
rating_Ly1×imm_Ly1	0.065165
eta×LEVclass_Ly2	0.065914
ben_FG_T0×LEVclass_Ly2	0.078429
sof_Lq8×LEVclass_Ly2	0.085678
X_1059_Ly1×ben_FG_T0	0.118018
PNneg_Ly1×LEVclass_Ly2	0.149409
eta	0.158293
ben_FG_T0×south	0.161868
X_1001_Ly1	0.162070
PNnull_Ly2	0.234353
rating_Ly2×no_aff	0.253314
rating_Ly2×south	0.347448
rating_Ly2×ben_FG_T0	0.374981
PNneg_Ly2×LEVclass_Ly2	0.650047
no_aff	0.899711

Table A7. Confusion matrices for each ML algorithm in the credit-constrained exercise

Panel A. Decision tree				
	$Y_{pred} = 0$	$Y_{pred} = 1$	Misclassification rate: 31.85%	
$Y_{actual} = 0$	8,408	23,100	TN: 26.68%	FN:10.64%
$Y_{actual} = 1$	6,558	55,033	FP: 73.3%	TP: 89.35%
Panel B. Random forest				
	$Y_{pred} = 0$	$Y_{pred} = 1$	Misclassification rate: 32.09%	
$Y_{actual} = 0$	10,243	21,265	TN: 32.5%	FN: 13.98%
$Y_{actual} = 1$	8,616	52,975	FP: 67.49%	TP: 86%
Panel C. LASSO regression				
	$Y_{pred} = 0$	$Y_{pred} = 1$	Misclassification rate: 33.83%	
$Y_{actual} = 0$	2,362	29,146	TN: 7.49%	FN: 3.82%
$Y_{actual} = 1$	2,354	59,237	FP: 92.5%	TP: 96.17%

Notes. Testing sample (2011). Y_{actual} is 1 if the actual status is to be credit constrained, 0 otherwise; Y_{pred} is 1 if a credit-constrained observation is predicted (predicted probability ≥ 0.5), 0 otherwise. FP is the *false positive rate* computed as the percentage of observations predicted positive, but that are actually negative, over the total number of actually negative observations; TP is the *true positive rate* computed as the percentage of observations predicted positive, that are actually positive, over the total number of actually positive observations; FN is the *false negative rate* computed as the percentage of observations predicted negative, but that are actually true, over the total number of actually positive observations; TN is the *true negative rate* computed as the percentage of observations predicted negative, but that are actually negative, over the total number of actually negative observations.

Table A8. Confusion matrices for each ML algorithm in the creditworthy exercise

Panel A. Decision tree				
	$Y_{pred} = 0$	$Y_{pred} = 1$	Misclassification rate: 20.02%	
$Y_{actual} = 0$	5,097	7,574	TN: 40.22%	FN: 13.75%
$Y_{actual} = 1$	11,066	69,362	FP: 59.77%	TP: 86.24%
Panel B. Random forest				
	$Y_{pred} = 0$	$Y_{pred} = 1$	Misclassification rate: 17.66%	
$Y_{actual} = 0$	4,948	7,723	TN:39.05%	FN: 10.84%
$Y_{actual} = 1$	8,726	71,702	FP: 60.95%	TP: 89.15%
Panel C. LASSO regression				
	$Y_{pred} = 0$	$Y_{pred} = 1$	Misclassification rate: 18.55%	
$Y_{actual} = 0$	3,661	9,010	TN: 28.89%	FN: 10.27%
$Y_{actual} = 1$	8,264	72,164	FP: 71.1%	TP: 89.72%

Notes. Testing sample (2011). Y_{actual} is 1 if the actual status is to be creditworthy, 0 otherwise; Y_{pred} is 1 if a creditworthy observation is predicted (predicted probability ≥ 0.5), 0 otherwise. FP is the *false positive rate* computed as the percentage of observations predicted positive, but that are actually negative, over the total number of actually negative observations; TP is the *true positive rate* computed as the percentage of observations predicted positive, that are actually positive, over the total number of actually positive observations; FN is the *false negative rate* computed as the percentage of observations predicted negative, but that are actually true, over the total number of actually positive observations; TN is the *true negative rate* computed as the percentage of observations predicted negative, but that are actually negative, over the total number of actually negative observations.

RECENTLY PUBLISHED "TEMI" (*)

N. 1183 – *Why do banks securitise their assets? Bank-level evidence from over one hundred countries in the pre-crisis period,* by Fabio Panetta and Alberto Franco Pozzolo.

N. 1184 – *Capital controls spillovers*, by Valerio Nispi Landi (July 2018).

N.1185 – *The macroeconomic effects of an open-ended asset purchase programme*, by Lorenzo Burlon, Alessandro Notarpietro and Massimiliano Pisani (July 2018).

N.1186 – *Fiscal buffers, private debt and recession: the good, the bad and the ugly*, by Nicoletta Batini, Giovanni Melina and Stefania Villa (July 2018).

N. 1187 – *Competition and the pass-through of unconventional monetary policy: evidence from TLTROs*, by Matteo Benetton and Davide Fantino (September 2018).

N. 1188 – *Raising aspirations and higher education: evidence from the UK's Widening Participation Policy*, by Lucia Rizzica (September 2018).

N. 1189 – *Nearly exact Bayesian estimation of non-linear no-arbitrage term structure models*, by Marcello Pericoli and Marco Taboga (September 2018).

N. 1190 – *Granular Sources of the Italian business cycle*, by Nicolò Gnocato and Concetta Rondinelli (September 2018).

N. 1191 – *Debt restructuring with multiple bank relationships*, by Angelo Baglioni, Luca Colombo and Paola Rossi (September 2018).

N. 1192 – *Exchange rate pass-through into euro area inflation. An estimated structural model*, by Lorenzo Burlon, Alessandro Notarpietro and Massimiliano Pisani (September 2018).

N. 1193 – *The effect of grants on university drop-out rates: evidence on the Italian case*, by Francesca Modena, Enrico Rettore and Giulia Martina Tanzi (September 2018).

N. 1194 – *Potential output and microeconomic heterogeneity*, by Davide Fantino (November 2018).

N. 1195 – *Immigrants, labor market dynamics and adjustment to shocks in the Euro Area*, by Gaetano Basso, Francesco D'Amuri and Giovanni Peri (November 2018).

N. 1196 – *Sovereign debt maturity structure and its costs*, by Flavia Corneli (November 2018).

N. 1197 – *Fiscal policy in the US: a new measure of uncertainty and its recent development*, by Alessio Anzuini and Luca Rossi (November 2018).

N. 1198 – *Macroeconomics determinants of the correlation between stocks and bonds*, by Marcello Pericoli (November 2018).

N. 1199 – *Bank capital constraints, lending supply and economic activity*, by Antonio M. Conti, Andrea Nobili and Federico M. Signoretti (November 2018).

N. 1200 – *The effectiveness of capital controls*, by Valerio Nispi Landi and Alessandro Schiavone (November 2018).

N. 1201 – *Contagion in the CoCos market? A case study of two stress events*, by Pierluigi Bologna, Arianna Miglietta and Anatoli Segura (November 2018).

N. 1202 – *Is ECB monetary policy more powerful during expansions?*, by Martina Cecioni (December 2018).

N. 1203 – *Firms' inflation expectations and investment plans*, by Adriana Grasso and Tiziano Ropele (December 2018).

N. 1204 – *Recent trends in economic activity and TFP in Italy with a focus on embodied technical progress*, by Alessandro Mistretta and Francesco Zollino (December 2018).

(*) Requests for copies should be sent to:
Banca d'Italia – Servizio Studi di struttura economica e finanziaria – Divisione Biblioteca e Archivio storico – Via Nazionale, 91 – 00184 Rome – (fax 0039 06 47922059). They are available on the Internet www.bancaditalia.it.

2017

AABERGE, R., F. BOURGUIGNON, A. BRANDOLINI, F. FERREIRA, J. GORNICK, J. HILLS, M. JÄNTTI, S. JENKINS, J. MICKLEWRIGHT, E. MARLIER, B. NOLAN, T. PIKETTY, W. RADERMACHER, T. SMEEDING, N. STERN, J. STIGLITZ, H. SUTHERLAND, *Tony Atkinson and his legacy*, Review of Income and Wealth, v. 63, 3, pp. 411-444, **WP 1138 (September 2017).**

ACCETTURO A., M. BUGAMELLI and A. LAMORGESE, *Law enforcement and political participation: Italy 1861-65*, Journal of Economic Behavior & Organization, v. 140, pp. 224-245, **WP 1124 (July 2017).**

ADAMOPOULOU A. and G.M. TANZI, *Academic dropout and the great recession*, Journal of Human Capital, V. 11, 1, pp. 35–71, **WP 970 (October 2014).**

ALBERTAZZI U., M. BOTTERO and G. SENE, *Information externalities in the credit market and the spell of credit rationing*, Journal of Financial Intermediation, v. 30, pp. 61–70, **WP 980 (November 2014).**

ALESSANDRI P. and H. MUMTAZ, *Financial indicators and density forecasts for US output and inflation*, Review of Economic Dynamics, v. 24, pp. 66-78, **WP 977 (November 2014).**

BARBIERI G., C. ROSSETTI and P. SESTITO, *Teacher motivation and student learning*, Politica economica/Journal of Economic Policy, v. 33, 1, pp.59-72, **WP 761 (June 2010).**

BENTIVOGLI C. and M. LITTERIO, *Foreign ownership and performance: evidence from a panel of Italian firms*, International Journal of the Economics of Business, v. 24, 3, pp. 251-273, **WP 1085 (October 2016).**

BRONZINI R. and A. D'IGNAZIO, *Bank internationalisation and firm exports: evidence from matched firm-bank data*, Review of International Economics, v. 25, 3, pp. 476-499 **WP 1055 (March 2016).**

BRUCHE M. and A. SEGURA, *Debt maturity and the liquidity of secondary debt markets*, Journal of Financial Economics, v. 124, 3, pp. 599-613, **WP 1049 (January 2016).**

BURLON L., *Public expenditure distribution, voting, and growth*, Journal of Public Economic Theory,, v. 19, 4, pp. 789–810, **WP 961 (April 2014).**

BURLON L., A. GERALI, A. NOTARPIETRO and M. PISANI, *Macroeconomic effectiveness of non-standard monetary policy and early exit. a model-based evaluation*, International Finance, v. 20, 2, pp.155-173, **WP 1074 (July 2016).**

BUSETTI F., *Quantile aggregation of density forecasts*, Oxford Bulletin of Economics and Statistics, v. 79, 4, pp. 495-512, **WP 979 (November 2014).**

CESARONI T. and S. IEZZI, *The predictive content of business survey indicators: evidence from SIGE*, Journal of Business Cycle Research, v.13, 1, pp 75–104, **WP 1031 (October 2015).**

CONTI P., D. MARELLA and A. NERI, *Statistical matching and uncertainty analysis in combining household income and expenditure data*, Statistical Methods & Applications, v. 26, 3, pp 485–505, **WP 1018 (July 2015).**

D'AMURI F., *Monitoring and disincentives in containing paid sick leave*, Labour Economics, v. 49, pp. 74-83, **WP 787 (January 2011).**

D'AMURI F. and J. MARCUCCI, *The predictive power of google searches in forecasting unemployment*, International Journal of Forecasting, v. 33, 4, pp. 801-816, **WP 891 (November 2012).**

DE BLASIO G. and S. POY, *The impact of local minimum wages on employment: evidence from Italy in the 1950s*, Journal of Regional Science, v. 57, 1, pp. 48-74, **WP 953 (March 2014).**

DEL GIOVANE P., A. NOBILI and F. M. SIGNORETTI, *Assessing the sources of credit supply tightening: was the sovereign debt crisis different from Lehman?*, International Journal of Central Banking, v. 13, 2, pp. 197-234, **WP 942 (November 2013).**

DEL PRETE S., M. PAGNINI, P. ROSSI and V. VACCA, *Lending organization and credit supply during the 2008–2009 crisis*, Economic Notes, v. 46, 2, pp. 207–236, **WP 1108 (April 2017).**

DELLE MONACHE D. and I. PETRELLA, *Adaptive models and heavy tails with an application to inflation forecasting*, International Journal of Forecasting, v. 33, 2, pp. 482-501, **WP 1052 (March 2016).**

FEDERICO S. and E. TOSTI, *Exporters and importers of services: firm-level evidence on Italy*, The World Economy, v. 40, 10, pp. 2078-2096, **WP 877 (September 2012).**

GIACOMELLI S. and C. MENON, *Does weak contract enforcement affect firm size? Evidence from the neighbour's court*, Journal of Economic Geography, v. 17, 6, pp. 1251-1282, **WP 898 (January 2013).**

LOBERTO M. and C. PERRICONE, *Does trend inflation make a difference?*, Economic Modelling, v. 61, pp. 351–375, **WP 1033 (October 2015).**

MANCINI A.L., C. MONFARDINI and S. PASQUA, *Is a good example the best sermon? Children's imitation of parental reading*, Review of Economics of the Household, v. 15, 3, pp 965–993, **D No. 958 (April 2014).**

MEEKS R., B. NELSON and P. ALESSANDRI, *Shadow banks and macroeconomic instability*, Journal of Money, Credit and Banking, v. 49, 7, pp. 1483–1516, **WP 939 (November 2013).**

MICUCCI G. and P. ROSSI, *Debt restructuring and the role of banks' organizational structure and lending technologies*, Journal of Financial Services Research, v. 51, 3, pp 339–361, **WP 763 (June 2010).**

MOCETTI S., M. PAGNINI and E. SETTE, *Information technology and banking organization*, Journal of Journal of Financial Services Research, v. 51, pp. 313-338, **WP 752 (March 2010).**

MOCETTI S. and E. VIVIANO, *Looking behind mortgage delinquencies*, Journal of Banking & Finance, v. 75, pp. 53-63, **WP 999 (January 2015).**

NOBILI A. and F. ZOLLINO, *A structural model for the housing and credit market in Italy*, Journal of Housing Economics, v. 36, pp. 73-87, **WP 887 (October 2012).**

PALAZZO F., *Search costs and the severity of adverse selection*, Research in Economics, v. 71, 1, pp. 171-197, **WP 1073 (July 2016).**

PATACCHINI E. and E. RAINONE, *Social ties and the demand for financial services*, Journal of Financial Services Research, v. 52, 1–2, pp 35–88, **WP 1115 (June 2017).**

PATACCHINI E., E. RAINONE and Y. ZENOU, *Heterogeneous peer effects in education*, Journal of Economic Behavior & Organization, v. 134, pp. 190–227, **WP 1048 (January 2016).**

SBRANA G., A. SILVESTRINI and F. VENDITTI, *Short-term inflation forecasting: the M.E.T.A. approach*, International Journal of Forecasting, v. 33, 4, pp. 1065-1081, **WP 1016 (June 2015).**

SEGURA A. and J. SUAREZ, *How excessive is banks' maturity transformation?*, Review of Financial Studies, v. 30, 10, pp. 3538–3580, **WP 1065 (April 2016).**

VACCA V., *An unexpected crisis? Looking at pricing effectiveness of heterogeneous banks*, Economic Notes, v. 46, 2, pp. 171–206, **WP 814 (July 2011).**

VERGARA CAFFARELI F., *One-way flow networks with decreasing returns to linking*, Dynamic Games and Applications, v. 7, 2, pp. 323-345, **WP 734 (November 2009).**

ZAGHINI A., *A Tale of fragmentation: corporate funding in the euro-area bond market*, International Review of Financial Analysis, v. 49, pp. 59-68, **WP 1104 (February 2017).**

2018

ADAMOPOULOU A. and E. KAYA, *Young adults living with their parents and the influence of peers*, Oxford Bulletin of Economics and Statistics,v. 80, pp. 689-713, **WP 1038 (November 2015).**

ANDINI M., E. CIANI, G. DE BLASIO, A. D'IGNAZIO and V. SILVESTRINI, *Targeting with machine learning: an application to a tax rebate program in Italy*, Journal of Economic Behavior & Organization, v. 156, pp. 86-102, **WP 1158 (December 2017).**

BARONE G., G. DE BLASIO and S. MOCETTI, *The real effects of credit crunch in the great recession: evidence from Italian provinces*, Regional Science and Urban Economics, v. 70, pp. 352-59, **WP 1057 (March 2016).**

BELOTTI F. and G. ILARDI *Consistent inference in fixed-effects stochastic frontier models*, Journal of Econometrics, v. 202, 2, pp. 161-177, **WP 1147 (October 2017).**

BERTON F., S. MOCETTI, A. PRESBITERO and M. RICHIARDI, *Banks, firms, and jobs*, Review of Financial Studies, v.31, 6, pp. 2113-2156, **WP 1097 (February 2017).**

BOFONDI M., L. CARPINELLI and E. SETTE, *Credit supply during a sovereign debt crisis*, Journal of the European Economic Association, v.16, 3, pp. 696-729, **WP 909 (April 2013).**

BOKAN N., A. GERALI, S. GOMES, P. JACQUINOT and M. PISANI, *EAGLE-FLI: a macroeconomic model of banking and financial interdependence in the euro area*, Economic Modelling, v. 69, C, pp. 249-280, **WP 1064 (April 2016).**

BRILLI Y. and M. TONELLO, *Does increasing compulsory education reduce or displace adolescent crime? New evidence from administrative and victimization data*, CESifo Economic Studies, v. 64, 1, pp. 15–4, **WP 1008 (April 2015).**

BUONO I. and S. FORMAI *The heterogeneous response of domestic sales and exports to bank credit shocks,* Journal of International Economics, v. 113, pp. 55-73, **WP 1066 (March 2018).**

BURLON L., A. GERALI, A. NOTARPIETRO and M. PISANI, *Non-standard monetary policy, asset prices and macroprudential policy in a monetary union,* Journal of International Money and Finance, v. 88, pp. 25-53, **WP 1089 (October 2016).**

CARTA F. and M. DE PHLIPPIS, *You've Come a long way, baby. Husbands' commuting time and family labour supply,* Regional Science and Urban Economics, v. 69, pp. 25-37, **WP 1003 (March 2015).**

CARTA F. and L. RIZZICA, *Early kindergarten, maternal labor supply and children's outcomes: evidence from Italy,* Journal of Public Economics, v. 158, pp. 79-102, **WP 1030 (October 2015).**

CASIRAGHI M., E. GAIOTTI, L. RODANO and A. SECCHI, *A "Reverse Robin Hood"? The distributional implications of non-standard monetary policy for Italian households,* Journal of International Money and Finance, v. 85, pp. 215-235, **WP 1077 (July 2016).**

CECCHETTI S., F. NATOLI and L. SIGALOTTI, *Tail co-movement in inflation expectations as an indicator of anchoring,* International Journal of Central Banking, v. 14, 1, pp. 35-71, **WP 1025 (July 2015).**

CIANI E. and C. DEIANA, *No Free lunch, buddy: housing transfers and informal care later in life,* Review of Economics of the Household, v.16, 4, pp. 971-1001, **WP 1117 (June 2017).**

CIPRIANI M., A. GUARINO, G. GUAZZAROTTI, F. TAGLIATI and S. FISHER, *Informational contagion in the laboratory,* Review of Finance, v. 22, 3, pp. 877-904, **WP 1063 (April 2016).**

DE BLASIO G, S. DE MITRI, S. D'IGNAZIO, P. FINALDI RUSSO and L. STOPPANI, *Public guarantees to SME borrowing. A RDD evaluation,* Journal of Banking & Finance, v. 96, pp. 73-86, **WP 1111 (April 2017).**

GERALI A., A. LOCARNO, A. NOTARPIETRO and M. PISANI, *The sovereign crisis and Italy's potential output,* Journal of Policy Modeling, v. 40, 2, pp. 418-433, **WP 1010 (June 2015).**

LIBERATI D., *An estimated DSGE model with search and matching frictions in the credit market,* International Journal of Monetary Economics and Finance (IJMEF), v. 11, 6, pp. 567-617, **WP 986 (November 2014).**

LINARELLO A., *Direct and indirect effects of trade liberalization: evidence from Chile,* Journal of Development Economics, v. 134, pp. 160-175, **WP 994 (December 2014).**

NUCCI F. and M. RIGGI, *Labor force participation, wage rigidities, and inflation,* Journal of Macroeconomics, v. 55, 3 pp. 274-292, **WP 1054 (March 2016).**

RIGON M. and F. ZANETTI, *Optimal monetary policy and fiscal policy interaction in a non_ricardian economy,* International Journal of Central Banking, v. 14 3, pp. 389-436, **WP 1155 (December 2017).**

SEGURA A., *Why did sponsor banks rescue their SIVs?,* Review of Finance, v. 22, 2, pp. 661-697, **WP 1100 (February 2017).**

2019

CIANI E. and P. FISHER, *Dif-in-dif estimators of multiplicative treatment effects,* Journal of Econometric Methods, v. 8. 1, pp. 1-10, **WP 985 (November 2014).**

FORTHCOMING

ACCETTURO A., W. DI GIACINTO, G. MICUCCI and M. PAGNINI, *Geography, productivity and trade: does selection explain why some locations are more productive than others?,* Journal of Regional Science, **WP 910 (April 2013).**

ALBANESE G., G. DE BLASIO and P. SESTITO, *Trust, risk and time preferences: evidence from survey data,* International Review of Economics, **WP 911 (April 2013).**

APRIGLIANO V., G. ARDIZZI and L. MONTEFORTE, *Using the payment system data to forecast the economic activity,* International Journal of Central Banking, **WP 1098 (February 2017).**

ARNAUDO D., G. MICUCCI, M. RIGON and P. ROSSI, *Should I stay or should I go? Firms' mobility across banks in the aftermath of the financial crisis,* Italian Economic Journal / Rivista italiana degli economisti, **WP 1086 (October 2016).**

BELOTTI F. and G. ILARDI, *Consistent inference in fixed-effects stochastic frontier models*, Journal of Econometrics, **WP 1147 (October 2017).**

BUSETTI F. and M. CAIVANO, *Low frequency drivers of the real interest rate: empirical evidence for advanced economies*, International Finance, **WP 1132 (September 2017).**

CHIADES P., L. GRECO, V. MENGOTTO, L. MORETTI and P. VALBONESI, *Fiscal consolidation by intergovernmental transfers cuts? Economic Modelling*, **WP 1076 (July 2016).**

CIANI E., F. DAVID and G. DE BLASIO, *Local responses to labor demand shocks: a re-assessment of the case of Italy*, IMF Economic Review, **WP 1112 (April 2017).**

COLETTA M., R. DE BONIS and S. PIERMATTEI, *Household debt in OECD countries: the role of supply-side and demand-side factors*, Social Indicators Research, **WP 989 (November 2014).**

CORSELLO F. and V. NISPI LANDI, *Labor market and financial shocks: a time-varying analysis*, Journal of Money, Credit and Banking, **WP 1179 (June 2018).**

COVA P., P. PAGANO and M. PISANI, *Domestic and international macroeconomic effects of the Eurosystem Expanded Asset Purchase Programme*, IMF Economic Review, **WP 1036 (October 2015).**

D'AMURI F., *Monitoring and disincentives in containing paid sick leave*, Labour Economics, **WP 787 (January 2011).**

D'IGNAZIO A. and C. MENON, *The causal effect of credit Guarantees for SMEs: evidence from Italy*, Scandinavian Journal of Economics, **WP 900 (February 2013).**

ERCOLANI V. and J. VALLE E AZEVEDO, *How can the government spending multiplier be small at the zero lower bound?*, Macroeconomic Dynamics, **WP 1174 (April 2018).**

FEDERICO S. and E. TOSTI, *Exporters and importers of services: firm-level evidence on Italy*, The World Economy, **WP 877 (September 2012).**

GERALI A. and S. NERI, *Natural rates across the Atlantic*, Journal of Macroeconomics, **WP 1140 (September 2017).**

GIACOMELLI S. and C. MENON, *Does weak contract enforcement affect firm size? Evidence from the neighbour's court*, Journal of Economic Geography, **WP 898 (January 2013).**

GIORDANO C., M. MARINUCCI and A. SILVESTRINI, *The macro determinants of firms' and households' investment: evidence from Italy*, Economic Modelling, **WP 1167 (March 2018).**

NATOLI F. and L. SIGALOTTI, *Tail co-movement in inflation expectations as an indicator of anchoring*, International Journal of Central Banking, **WP 1025 (July 2015).**

RIGGI M., *Capital destruction, jobless recoveries, and the discipline device role of unemployment*, Macroeconomic Dynamics, **WP 871 (July 2012).**

RIZZICA L., *Raising aspirations and higher education. evidence from the UK's widening participation policy*, Journal of Labor Economics, **WP 1188 (September 2018).**

SEGURA A., *Why did sponsor banks rescue their SIVs?*, Review of Finance, **WP 1100 (February 2017).**