Machine Learning in the Service of Policy Targeting: the Case of Public Credit Guarantees

M. Andini†, E. Ciani†, G. de Blasio†,
A. D'Ignazio†, M. Boldrini‡ , A. Paladini§

†Bank of Italy, ‡University of Bologna, §University of Rome La Sapienza

Modelling with Big Data and Machine Learning
26-27 November 2018

Prediction policy problems

- The effectiveness of a public policy depends (also) on who benefits from it
- Predicting those for whom the policy is most likely to be effective is a **prediction policy problem** (Kleinberg et al., 2017)
- The idea is to provide the policy maker with a decision rule to identify these agents
- Machine Learning (ML) algorithms are suited for this prediction task
- But **via prediction we may not obtain what we want**
- We need to assess whether targeting-on-prediction improves effectiveness by using tools from **causal inference**

Public Guarantee Schemes

- In this paper we focus on the Italian public credit guarantee scheme
- In principle, Credit guarantee programs should target firms that are both **credit-worthy** (for financial stability of the scheme) and **credit-rationed** (to increase additionality)
- In practice, the current rule (also in other countries) is only based on firms' financial soundness
- We use ML to train (estimate) two predition models to identify, on the basis of observables, which firms are **both** credit-rationed and credit-worthy
- We **contrast** a targeting rule based on these predictions with the actual allocation rule and we **evaluate** wheter effectiveness would increase

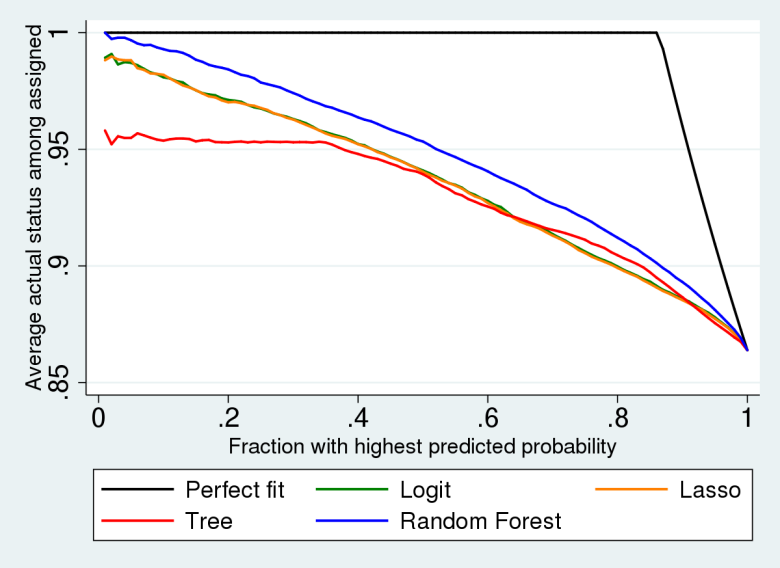
Empirical strategy: ML modeling

- We focus on **loan applications** in 2011-12 as recorded in the *prima informazione* system. We keep only those with Credit register (firm-bank) and Cerved (balance sheet) data
- We follow two steps:

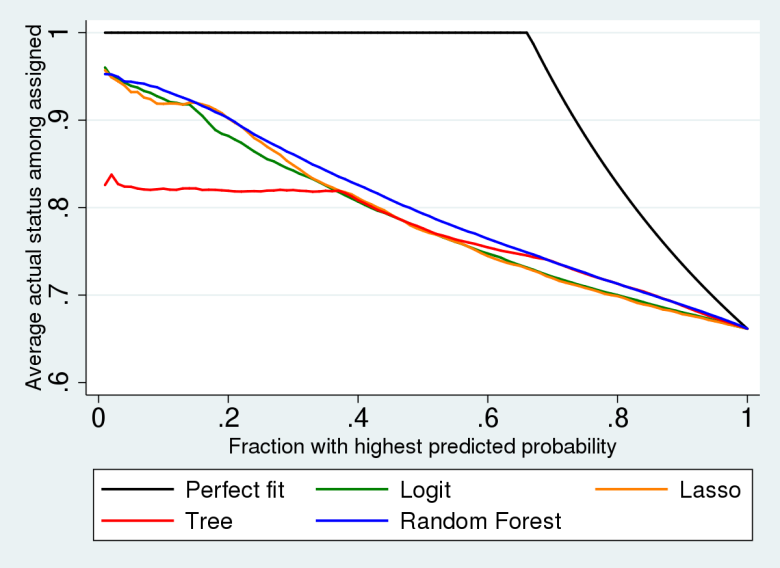
1. Using the 2011 sample ($\approx$ 280,000 obs) we train and test **two separate ML algorithms**
   - **credit-constrained** (observed variation in bank loans, as in Jiménez et al, AER, 2012)
   - **credit-worthy** (observed *sofferenze rettificate*)

   using **108 firm-level variables** as predictors and comparing three off-the-shelf ML algorithms: decision TREE, logistic LASSO and RANDOM FOREST
2. On new data (the 2012 sample) we evaluate whether effectiveness could be increased by targeting firms that are predicted to be both credit-worthy and credit-constrained (ML target)

Prediction accuracy for credit-worthy (2011, test sample)



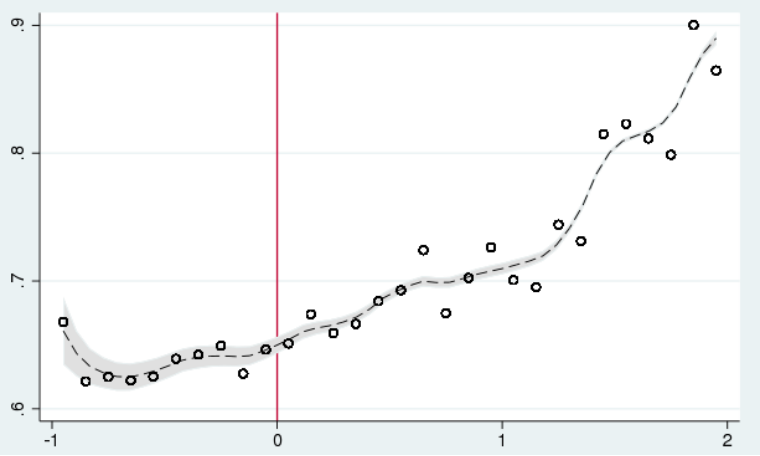
Prediction accuracy for credit-rationed (2011, test sample)



ML rule vs the actual Guarantee fund rule

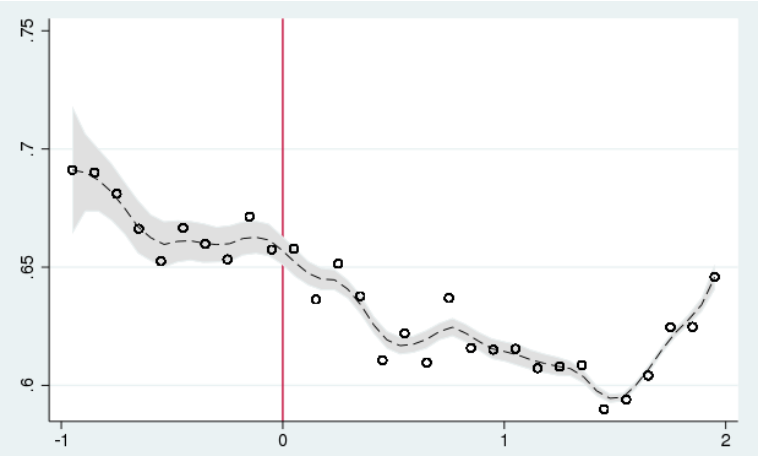
On the 2012 sample we compare predictions and actual GF eligibility

Figure: Predicted credit-worthy (y) vs Guarantee fund score (x)



ML rule vs the actual Guarantee fund rule (cont.)

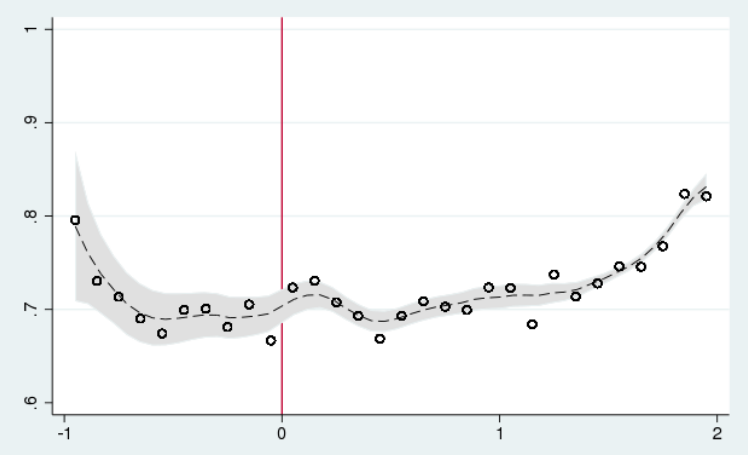
Figure: Predicted rationed (y) vs Guarantee fund score (x)



ML rule vs the actual Guarantee fund rule (cont.)

We define as ML targets those for which the probability of both status is ≥ 0.5

Figure: ML target (y) vs Guarantee fund score (x)



Impact of the alternative ML rule on GF's effectiveness

- We exploit data on Guarantee fund beneficiary firms and run on the 2012 sample 3 exercises:
  1. Compare various observed outcomes for treated (beneficiary) ML targeted firms vs treated-non ML targeted firms (Kleinberg et al's contraction)
  2. Compare counterfactual outcomes for non-treated ML targeted firms vs observed outcomes of treated-non ML targeted firms (Kleinberg et al's re-targeting)
  3. Relax the selection on observables assumption and run a fuzzy-RDD by splitting the sample into ML target vs ML non target

Impact of the alternative ML rule on GF's effectiveness/1

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

	ML target =1	ML target =0	Diff.	p-value
granted loans 2011-15	0.08	-0.48	0.56	0.00
adjusted bad loans 2015	0.03	0.06	-0.03	0.00
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. MLtarget=1: average performance in the period 2011-2015 for the set of firms that received the Fund guarantee and are targeted by our ML algorithm (combined prediction from random forest). MLtarget=0: average performance in the period 2011-2015 for the set of firms that received the Fund guarantee and are not targeted by our ML algorithm. Growth rates are computed as logarithm changes.

Impact of the alternative ML rule on GF's effectiveness/2

Table: Re-ranking

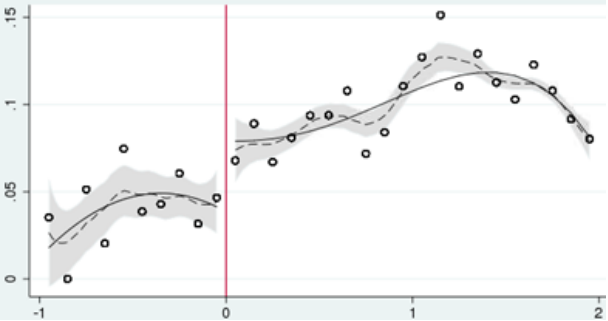
	Re-ranking =1	Re-ranking =0	Diff.	p-value
granted loans 2011-15	0.18	-0.48	0.66	0.00
adjusted bad loans 2015	0.03	0.06	-0.02	0.00
observations	1,303	2,869		

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 2011-2015 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 3504 firms that are targeted by ML but not eligible according to the Fund rules (Table 5); for 1303 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 report the average performance of the matched-treated firms. Re-ranking=0: average performance in the period 2011-2015 for the subset of firms that received the Fund guarantee and are not targeted by our ML algorithm. Growth rates are computed as logarithm changes.

Impact of the alternative ML rule on GF's effectiveness/3

- We exploit the current design to run an RDD regression separately by ML target status
- We check that first stage and balancing (also for predicted probabilities) is respected in both subsamples

Figure: Guarantee recipient (y) vs fund score (x)



Impact of the alternative ML rule on GF's effect./3 (cont.)

Table: Fuzzy-RDD analysis, non-parametric estimates, Wald estimates

	Full sample	ML target =1	ML target =0
granted loans 2011-15	0.652*	0.794*	0.433
	(0.363)	(0.468)	(0.467)
adjusted bad loans 2015	-0.089	-0.002	-0.139
	(0.500)	(0.754)	(0.647)

Notes. *** p-val$\leq$0.01, ** p-val$\leq$0.05, * p-val$\leq$0.1. Selected sample of 62,994 firms (see Section 5.2). Fuzzy-RDD non parametric estimates. The optimal bandwidth has been retrieved by Imbens and Kalyanaraman (2009) procedure. Outliers below 5 or above 95 percentile were dropped. Standard errors in brackets.

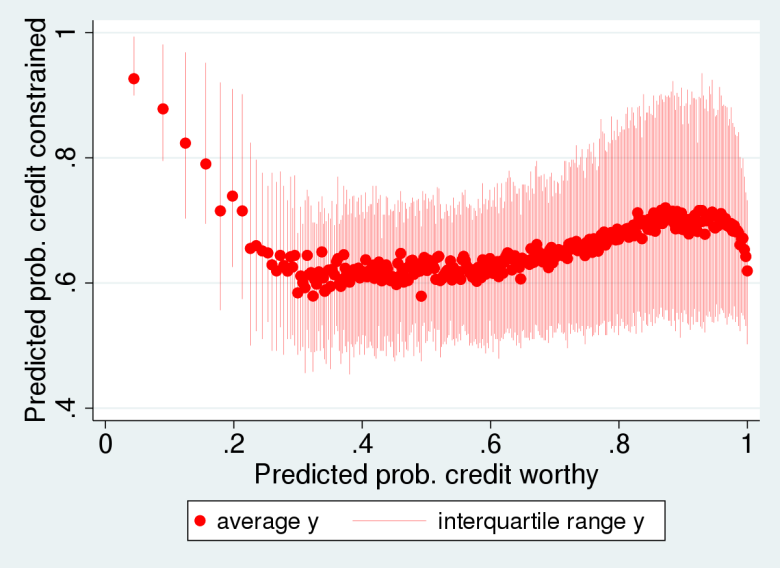
Issues

- Omitted payoffs (Kleinberg et al, 2017)
  - Any targeting rule might end up having other effects
  - We check two relevant issues: the distribution across regions and across different banks
  - The GF eligibility tends to favor of firms whose main reference bank belongs to a group and that are in the Centre-North
  - Conversely, ML eligibility is not correlated with banks belonging to a group and is in favor of firms from the South
- Transparency
  - Our favorite ML prediction model (random forest) is kind of a black box
  - This might be a concern
  - A decision tree provides more transparency, but we lose quite a lot in accuracy
  - Formal vs actual transparency: targeting via ML forces us to clearly indicate the purpose and assess whether the rule is fit for it

Issues (cont.)

- Contamination
  - The dataset is "contaminated" as it also contains firms that already receive the Fund guarantee
  - We also re-estimated the algorithm for rationed firms excluding the guarantee recipients, but they are few and predictions are similar
  - This issue boils down to the question about gains from ML targeting: if contamination annihilates the predictive power, then we should find no gains

Conclusions

- We discuss how ML algorithms can be used to re-target a public credit guarantee scheme
- We focused on the issue relative to the selection of firms that are not only credit-worthy, but also rationed
- Our approach tries to combine the pure ML-predictive part with a validation exercise based on counterfactual methods: prediction accuracy is not enough!
- The current reform of the Guarantee scheme goes in the direction of favoring the access of rationed firms, but to do so it still looks at a single dimension - the probability of default - assessed through a black box (and not fully documented) model

Appendix: Predicted rationed vs predicted worthy