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

M. Andini E. Ciani A. D’Ignazio G. de Blasio (Bank of Italy)
M. Boldrini (University of Bologna)
A. Paladini (University of Rome La Sapienza)
ML for policy targeting

• The effectiveness of a public policy depends (also) on who benefits from it (targeting). ML algorithms can be used to predict those for whom the policy is most likely to be effective (providing the policy maker with a rule to assign a treatment)

• Application: Italian public credit guarantee scheme (FG). 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)

• We use ML (d-Tree, Lasso, RF) to predict firms that are both credit-worthy (y=no NPL) and credit-rationed (y=no credit after a PI request) and we contrast the ML targeting rule with the actual allocation rule

• Results: 20% of the firms that are FG eligible are discarded by ML; 75% of the FG non eligible are retrieved by ML → the actual allocation places too much weight on firm creditworthiness, neglecting credit rationing
Ex-post evidence and issues

- We exploit data on FG beneficiary and run 3 exercises: 1) compare observed outcomes for treated (beneficiary) ML targeted firms vs treated-non ML targeted firms; 2) compare counterfactual outcomes for non-treated ML-targeted firms vs observed outcomes of treated non-ML-targeted firms; 3) elaborate on the actual allocation rule and estimate a F-RDD by target status

- **Huge gains in terms of effectiveness** (the resources granted to firms that are not targeted by ML reach 46.5%) but many issues (information requirements, contamination, etc):
  - RF is kind of a **black box**. A d-Tree provides more transparency, but we lose quite a lot in accuracy
  - **Formal vs substantive transparency**: targeting via ML forces us to clearly indicate the policy objective and assess whether the rule is fit for it
  - **Additional policy objectives**: the distribution across regions and across different banks