Journal of Economic Behavior and Organization 156 (2018) 86–102

Contents lists available at ScienceDirect

Journal of Economic Behavior and Organization

journal homepage: www.elsevier.com/locate/jebo

Targeting with machine learning: An application to a tax rebate program in Italy☆

Monica Andini[a], Emanuele Ciani[a,b], Guido de Blasio[a], Alessio D'Ignazio[a,*], Viola Salvestrini[c]

[a] Bank of Italy, Structural Economic Analysis Directorate, Via Nazionale 91, Roma 00184, Italy
[b] Centre for the Analysis of Public Policies, University of Modena and Reggio Emilia, Viale Berengario 51, Modena 41121, Italy
[c] London School of Economics and Political Science, Department of Economics, Houghton Street, London WC2A 2AE, United Kingdom

ARTICLE INFO

Article history:
Received 22 September 2017
Revised 23 July 2018
Accepted 11 September 2018
Available online 25 October 2018

JEL Codes:
C5
H3

Keywords:
Machine learning
Prediction
Program evaluation
Fiscal stimulus

ABSTRACT

This paper shows how machine learning (ML) methods can be used to improve the effectiveness of public schemes and inform policy decisions. Focusing on a massive tax rebate scheme introduced in Italy in 2014, it shows that the effectiveness of the program would have significantly increased if the beneficiaries had been selected according to a transparent and easily interpretable ML algorithm. Then, some issues in estimating and using ML for the actual implementation of public policies, such as transparency and accountability, are critically discussed.

1. Introduction

Machine Learning (ML) algorithms have been developed in the computer science and statistical literature. Differently from the econometrics literature, which mainly points towards reducing the estimator bias, ML algorithms focus on minimizing the out-of-sample prediction error (Athey and Imbens, 2017; Mullainathan and Spiess, 2017). Such algorithms are gaining popularity among economists, providing them with a new toolbox which is useful to deal with purely predictive tasks (Varian, 2014). In particular, a stream of research has focused on the so-called "prediction policy problems", a term coined by Kleinberg et al. (2015). Early applications include: (i) predicting the riskiest patients for which a joint replacement would be futile (Kleinberg et al., 2015); (ii) improving over judges' decision on whether to detain or release arrestees

☆ This paper was partly written while Viola Salvestrini was an intern at the Structural Economic Analysis Directorate of the Bank of Italy. For useful comments and suggestions, the authors would like to thank the editor Scott Adams, two anonymous referees, Audinga Baltrunaite, Stefano Gagliarducci, Christian Hansen, Matthew Harding, Giuseppe Ilardi, Andrea Neri, Enrico Rettore, Paolo Sestito and the participants at the *IX ESPAnet Conference* (Macerata, September 2016), the *First Dondena Workshop on Public Policy* (Milan, November 2016), the *Bank of Italy Seminar Series* (Rome, November 2016), the *Royal Economic Society Annual Conference* (Bristol, April 2017), the *Bank of Italy Internal Workshop on "Big Data"* (Rome, June 2017), the *Bank of Italy Workshop on "Harnessing Big Data & Machine Learning Technologies for Central Banks"* (Rome, March 2018). The views expressed in this paper are those of the authors and do not necessarily reflect those of the institutions they are affiliated with.

* Corresponding author.
E-mail address: alessio.dignazio@bancaditalia.it (A. D'Ignazio).

https://doi.org/10.1016/j.jebo.2018.09.010
0167-2681/

as they await adjudication of their case (Kleinberg et al., 2018); (iii) targeting restaurant hygiene inspections (Kang et al., 2013); (iv) predicting highest risk youth for anti-violence interventions (Chandler et al., 2011); (v) predicting the effectiveness of teachers in terms of value added (Rockoff et al., 2011); (vi) hiring police officers who will not behave violently, as well as promoting the best teachers only (Chalfin et al., 2016); (vii) improve poverty targeting (McBride and Nichols, 2015).

We extend this literature by focusing on a tax rebate scheme introduced in Italy in 2014, with the main purpose of boosting household consumption. The program, named the "80 euro" bonus,[1] has been a centerpiece of the Italian government policy efforts to counterbalance the negative consequences of the Great Recession.[2] We aim at understanding how ML methods can help improving the targeting of the scheme. In principle, those who are treated under a program might have different propensities to put in practice the behavior that the policy maker wants to incentivize, or have different payoffs from a given treatment. ML algorithms can help targeting the program towards those who are most likely to behave in the desired way, or have the highest payoff, so to increase the overall policy effectiveness. Specifically, we contrast the group targeted by the actual scheme - employees with a gross annual income between €8,145 and €26,000 - with the group that would be selected by a ML algorithm designed to increase the impact of the program on consumption. Using the Bank of Italy's Survey on Household Income and Wealth (SHIW), we start by showing that reporting difficulties in making ends meet - which we use as a proxy for being consumption constrained - is associated with a larger effect of the bonus on consumption. We therefore use this variable, observable only in the survey, as outcome in our targeting exercise. We assume to be in the position of the Italian government when the scheme was designed (early 2014), and to have the task to tailor the allocation rule towards those households that are more likely to increase their consumption upon receiving the bonus. To accomplish this task, we use the 2010 and 2012 SHIW waves to build a ML prediction model that identifies, on the basis of observable variables, the households most likely to be consumption constrained. We then apply the model on the 2014 SHIW wave to predict the consumption constrained households (i.e., the ML targeted households) and compare this group with the actual recipients of the tax rebate.

We consider three popular and "off-the-shelf" (Athey, 2018) ML algorithms - decision tree, k-Nearest Neighbor clustering and random forest (see Hastie et al., 2009) - and one non ML method, the linear probability model. Although in our case all the ML algorithms perform similarly to the linear probability model in terms of prediction accuracy (probably because of the limited size of our dataset), one of them - the decision tree - has the important advantage of providing a simple and transparent decision rule that is based on few variables. Hence, our analysis is mostly performed using this algorithm, as it provides both a good prediction accuracy and an easily interpretable assignment mechanism, which can be transparently communicated to the general public by an accountable policy maker. We show that the gains obtained by using the decision tree targeting rule can be substantial: 29.5% of the actual expenditure (about 2 billion euro, yearly) has been allocated to recipients that are not identified as consumption constrained. Using our algorithm, a policy maker could improve the current allocation scheme in two ways: (i) by cutting the bonus channeled to recipients that are not ML targeted, which would result in the same impact on consumption as the current policy, but with a lower expenditure; (ii) by shifting the bonus from the households that are not ML targeted to the targeted ones, which would lead to a greater impact on consumption and keep the level of expenditure unchanged. We also discuss how to use our insights to design an effective allocation of the bonus for the case of pensioners, that have been proposed as another potential target group in the debate right after the introduction of the 80 euro bonus. Furthermore, policy makers may have other payoffs that need to be taken into account when designing an allocation scheme. These are what Kleinberg et al. (2018) call "omitted payoffs", which we discuss with reference to redistribution and labor supply. Finally, we discuss the issues of transparency and accountability, which lead to major concerns when ML is applied for policy decisions (see, for instance, Athey, 2017, 2018; as well as OECD, 2017), and those related to the information needed to implement the ML targeting rule.

The remainder of the paper proceeds as follows. Section 2 provides the relevant institutional details on the 80 euro scheme. Section 3 sketches the SHIW features and gathers information on how the tax credit is detected in the 2014 wave. Section 4 contains some descriptive evidence on the effect of the income tax credit on consumption. Section 5 describes the ML empirical framework used for our targeting analysis and presents the results. Section 6 discusses the consumption constrained definition, the extension to retirees, and the concerns in using ML for policy choice. Section 7 concludes.

2. The bonus: institutional details

During the second dip of the economic crisis, total consumption of Italian households dropped from nearly 970 billion euro in 2010 to 909 in 2013. While foreign demand kept supporting exports, the recession was prolonged by the stagnation of internal demand. A government crisis led to the appointment of a new Prime Minister in February 2014. One of the first announcements was the proposal of a new monetary transfer to households, aimed at boosting consumption. This proposal was reiterated on several occasions during the early weeks of the new government, and formally announced to the press as a "80 euro" transfer on the 12th of March. Fig. 1, which describes Google trends for search and news about "80 euro", clearly shows how the press and general public became fast aware of the bonus. The transfer was finally implemented as a tax credit with the Decree Law 66/2014, which was approved by the government on the 24th of April, and later ratified by

[1] Hereafter, the expressions *tax credit*, *tax rebate*, and *bonus* are used interchangeably.

[2] Our work is related to the papers studying the effect of the 80 euro bonus on consumption. See: Neri et al. (2017), Gagliarducci and Guiso (2015), Pinotti (2015).

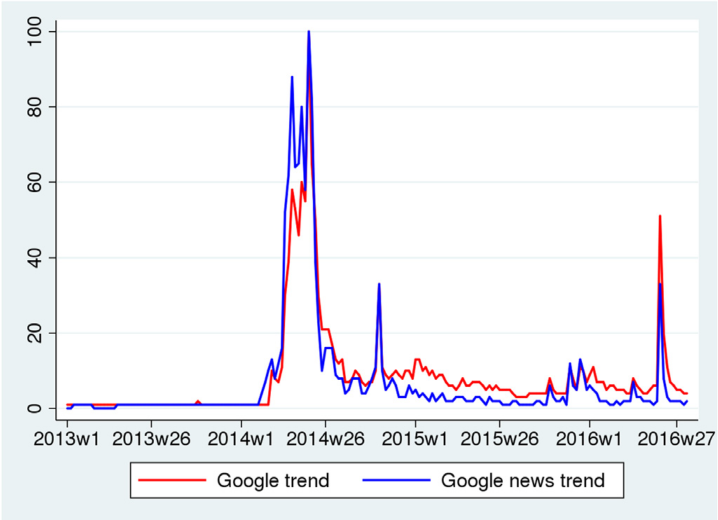


Fig. 1. Google trends for "80 euro", Italy
Notes. Search indexes on the vertical axis: 100 indicates the highest search frequency of the term over the period indicated in the horizontal axis; 0 indicates that no sufficient data were found for the term.

the Parliament. There was no debate about this bonus earlier than the change of government in 2014, hence our analysis should not be affected by possible anticipation effects. The scheme was financially considerable: according to government's estimates (Ministry of Economics and Finance, 2015), it entailed a transfer of almost 7 billion euro in 2014, equivalent to 0.4% of Italian GDP.

The benefit was designed as a tax credit. It targets employees and holders of similar income[3] with gross annual income between €8,145 and €26,000.[4] In particular, the tax credit amounts to €640 per year (i.e. €80 each month the policy was in place in 2014) if the gross annual income ranges between €8,145 and €24,000; for earnings between €24,000 and €26,000, the amount of the benefit is calculated as follows: $80 \times (26,000 - income)/2,000$. The tax credit is acknowledged automatically by the employer on the monthly salary paid without a specific request by the beneficiary, on the basis of the predicted annual income according to the current contract and as long as the gross (predicted) tax is greater than the tax deduction for employee income. An individual may receive the bonus on the basis of this predicted income, but will have to give it back if the actual income at the end of the year is outside the eligible range.[5] Several factors contributed to determine the allocation rule.[6] First of all, employees were chosen as recipients so that the tax credit could be automatically acknowledged by the withholding agent (usually the employer). This was clearly aimed at making the implementation of the program much easier and faster. Following the same rationale, the "*incapienti*" (namely, those who earn so little that they do not pay taxes) were left out, so to avoid the withholding agent to pay the transfer out of pocket.

3. The bonus in the SHIW

The Survey on Household Income and Wealth (SHIW) is conducted on a biennial basis by the Bank of Italy, to gather information on the economic behavior of the Italian families at the microeconomic level. The 2014 wave contains some specific questions on the income tax credit: households were asked if they received the bonus, how many beneficiaries were present within the family and the overall amount received. The overall size of the sample for the 2014 wave is 8,156 households. Since a necessary condition to be eligible for the bonus is to be an employee, we consider only households with at least one employee, for a total of 3,646 observations.[7] Given that the data were collected between January and July 2015, when the income referring to the previous year was secured, we can reasonably assume that individuals knew at the moment of the interview if they were entitled to the bonus or they had to give it back. We also make use of previous SHIW waves to build the forecasting model aimed at improving the targeting of the bonus. In order to keep information as homogeneous as possible, we only use the two waves collected after the beginning of the recession (that is, 2010 and 2012).

[3] For instance, freelancers, priests, cooperative workers, recipients of unemployment or disability benefits, recipients of scholarships and unemployment insurance (*Cassa Integrazione*).

[4] The bonus was granted under the condition of a net tax greater than zero. People with gross annual income below €8,145 do not pay taxes and were excluded. The threshold of €8,145 applies to individuals who worked the entire year, but it might be lower for workers who have been employed for less than 365 days.

[5] This will happen at the moment of the annual tax return, which has to be submitted between April and the beginning of July 2015 for the 2014 annual income.

[6] For a full description of the policy from a public finance point of view, see Baldini et al. (2015).

[7] We do not consider the other categories of bonus recipients, which however represent a small fraction of the beneficiaries.

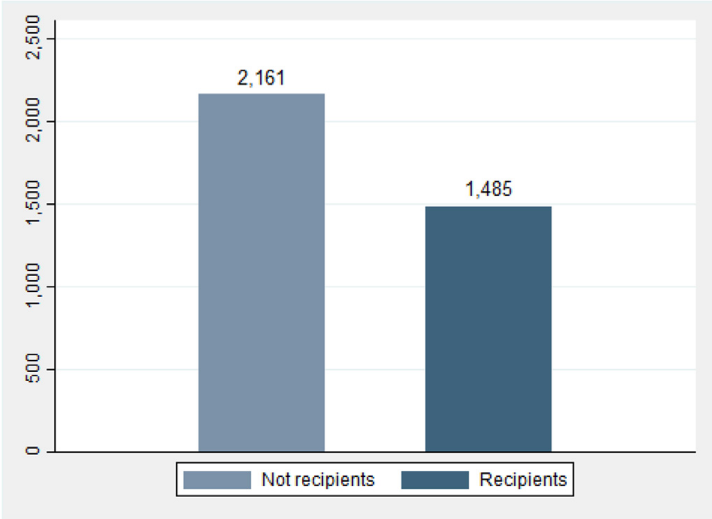


Fig. 2. Bonus recipients and not recipients
Notes. Elaboration on the SHIW 2014 data; households with at least one employee.

Although the dataset provides us with a large set of covariates, the actual income variable that determines the eligibility for the bonus is not observed. The survey collects only income net of taxes and social contributions, while the bonus was determined on the basis of gross income. It is not possible to simply invert the tax formula because it depends on a full set of deductions that are household specific (and not reported in the survey).

The SHIW collects both annual expenditure on durable goods and the average monthly non-durable consumption during the year. Households are also asked to report average monthly expenditure on food consumed at home and, separately, on food consumed outside home. The main limitation of these questions is that they are retrospective.[8] However, there are no alternative data. SHIW is the only data source for Italian households that contains all the information we need for our analysis: the receipt of the bonus, expenditure for consumption, income, wealth, demographic characteristics and, as detailed in Section 4, proxies for being consumption constrained.

4. Descriptive evidence

Using the 2014 SHIW wave, we first show that the households who received the bonus are not always the consumption constrained ones, and this signals serious problems of targeting. SHIW contains different proxies for being consumption constrained. The broader one is whether the household respondents reported that, with the available income, they have at least some difficulties to make ends meet.[9] Nearly 40% of the sample received the bonus (Fig. 2), but Fig. 3 shows that there does not seem to be a relevant difference in the incidence of difficulties in making ends meet between beneficiaries and non-beneficiaries. Another way to measure consumption constraints is to look at those who report to be liquidity constrained, i.e. the household has been at least partially rejected a request for a mortgage, or would have liked to apply for it but had not because of the assumption to be rejected. Among the recipients (Fig. 4), nearly 6% reported to be liquidity constrained, against 2.6% among the not recipients. Finally, a more standard way to look at consumption constraints is simply to look at households with low income. Fig. 5 splits bonus recipients and not recipients by income quartiles. Among the recipients, nearly 16% are in the top income quartile. This is due to the fact that we are considering household income rather than individual income.

We expect the consumption constrained households to consume more out of the bonus, as they report that their current income prevents them from making ends meet or that they have problems in accessing credit. To provide evidence supporting this expectation, we recover an estimate of the impact of the bonus on consumption by relying on selection on observables. As the bonus was automatically distributed to all eligible individuals on the basis of their tax-relevant information, it is reasonable to assume that no self-selection occurred. Having a higher external validity, as it is the case for selection on observables, is desirable for our targeting analysis that associates the consumption constrained status with household observables. Our estimates, as shown below in this section, are nevertheless in line with those provided by Neri et al. (2017),

[8] The Survey on Household Consumption, carried out by the National Statistical Institute and used by Gagliarducci and Guiso (2015) and Pinotti (2015), is based - instead - on diaries filled in by a sample of households, who are asked to report detailed expenditures during a single month. Although this reduces the risk of misreporting, the fact that consumption refers to a single month of the year increases the volatility of the measure, thus reducing the ability to detect the effect of the bonus. In this respect, the nature of the consumption variables in SHIW, which are referred to the whole year (although reported as totals or monthly averages), may be more adequate for our purpose.

[9] The question is “The income available to your family allows you to make ends meet with: (i) great difficulty; (ii) difficulty; (iii) some difficulty; (iv) fairly easily; (v) easily; (vi) very easily”.

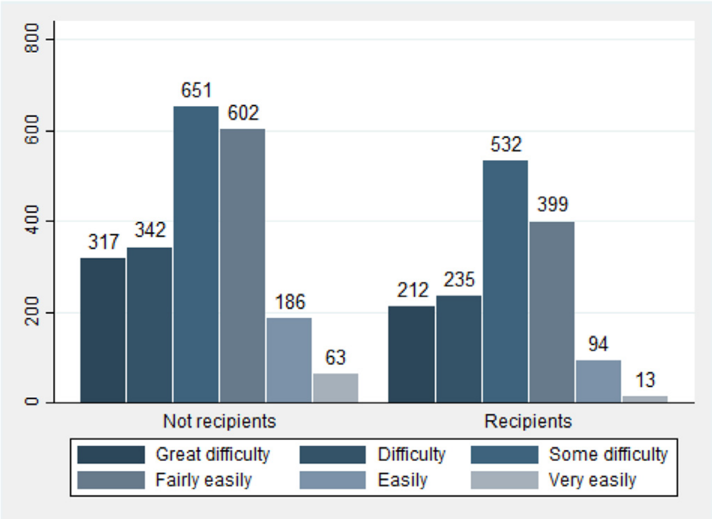


Fig. 3. Difficulty in making ends meet for bonus recipients and not recipients
Notes. Elaboration on the SHIW 2014 dataset; households with at least one employee.

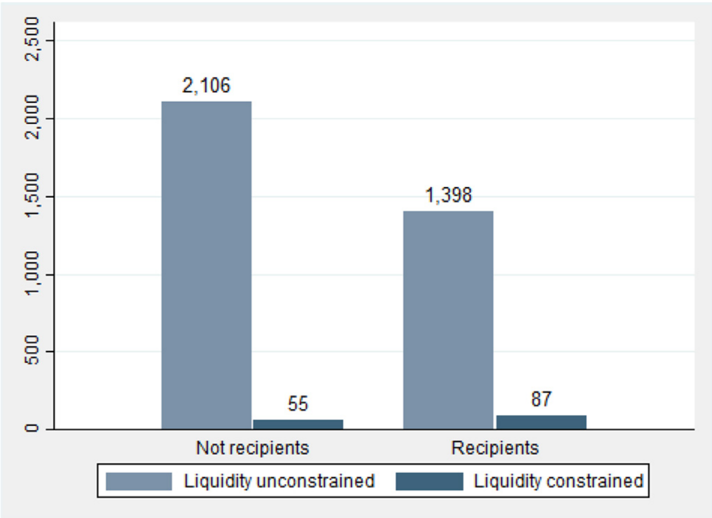


Fig. 4. Liquidity constraints for bonus recipients and not recipients
Notes. Elaboration on the SHIW 2014 dataset; households with at least one employee.

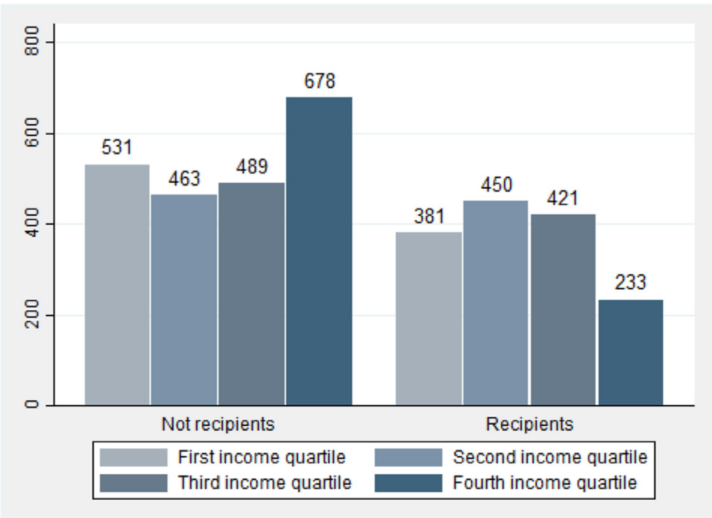


Fig. 5. Income quartiles for bonus recipients and not recipients
Notes. Elaboration on the SHIW 2014 dataset; households with at least one employee.

Table 1
Effect of the bonus on total non-durable consumption.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
bonus_amount	0.537*	0.114	0.134	0.217	0.214	0.208	0.373
	(0.281)	(0.273)	(0.275)	(0.275)	(0.276)	(0.276)	(0.276)
Income	0.023***	0.022***	0.022***	0.020***	0.020***	0.020***	0.020***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
ncomp		159.4***	157.5***	176.8***	177.1***	159.3***	153.5***
		(27.04)	(26.74)	(27.13)	(27.15)	(28.92)	(28.45)
ncomp2		−10.20**	−9.796**	−11.73***	−11.78***	−9.947**	−8.189*
		(4.379)	(4.335)	(4.303)	(4.309)	(4.381)	(4.200)
Age			0.797	2.042**	2.024**	1.972**	2.600***
			(0.764)	(0.866)	(0.866)	(0.862)	(0.894)
Diploma				69.53***	69.59***	69.73***	68.24***
				(21.14)	(21.14)	(21.13)	(20.96)
Degree				150.7***	151.5***	152.0***	148.7***
				(35.21)	(35.27)	(35.30)	(35.35)
Male					6.974	2.076	9.219
					(18.58)	(18.93)	(19.06)
Married						34.39	42.03*
						(22.45)	(22.35)
Constant	551.4***	247.1***	211.5***	94.08*	91.90*	109.4**	145.8**
	(27.72)	(34.22)	(44.67)	(55.16)	(55.26)	(55.47)	(74.23)
Regional FE	NO	NO	NO	NO	NO	NO	YES
Observations	3,646	3,646	3,646	3,646	3,646	3,646	3,646
R2	0.470	0.493	0.493	0.497	0.497	0.497	0.510

Notes. Estimation on the SHIW 2014 dataset. *-**-*** denotes statistical significance at 10%–5%–1%. Robust standard errors in parentheses.

who use a more rigorous identification approach. In particular, they use a difference-in-differences approach with the longitudinal component of the SHIW survey and suggest that households receiving the bonus spent around 50–60% of it. We cannot restrict our focus to the SHIW longitudinal component because its size would be too small for a reliable targeting analysis. Using different data (the Survey on Household Consumption matched with administrative data on labor income), Gagliarducci and Guiso (2015) apply a regression discontinuity design exploiting the fact that the bonus depends on labor income thresholds, and find a positive impact of the bonus on food and on mortgage payments, suggesting that the entire bonus goes to consumption. Using similar data as those of Gagliarducci and Guiso (2015), Pinotti (2015) argues that the effect on total consumption cannot be estimated with precision, and it may even be zero. We depart from studies that provide more internally valid estimates as the latter would be "local", thus making it more difficult to pursue our aim, i.e. obtain predictions on the overall sample.

Let $bonus\_amount_i$ be the average monthly amount of the bonus (in euro) received by household i. The main outcome of interest is the average monthly expenditure on consumption c_i (measured in euro) while x_i is a vector of household characteristics, including household income, household size (and its square) and several other characteristics, plus a constant. We then estimate

$$c_i = \delta bonus\_amount_i + \beta_x x_i + \varepsilon_i \tag{1}$$

$$E[\varepsilon_i | bonus\_amount_i, x_i] = 0 \tag{2}$$

Given that both c_i and $bonus\_amount_i$ are in euro, δ can be interpreted as the fraction of the bonus spent on consumption, conditional on x_i. We start by estimating the effect of the bonus on average monthly total non-durable consumption (Table 1; Supplementary Table A.1 provides a description of the variables used in the baseline specifications). We focus only on non-durable consumption because, as discussed by Neri et al. (2017), durable consumption tends to be more volatile and therefore it is difficult to detect the impact of the bonus. The total consumption of non-durables excludes rents (imputed or actual), mortgages, and in-kind benefits from the employer. To begin with, the average monthly amount of the bonus received and household annual disposable income (net of the bonus) are considered as regressors (column (1)). Since current consumption depends - possibly not linearly - on the size of the family, we also control for the number of components and its square (column (2)). Then, we subsequently add a rich set of demographic characteristics such as age, education, gender, and marital status of the household head (columns (3)–(6)). In column (7), regional fixed effects are included to capture specific factors that might affect all the people residing in the same area. Then, in line with the debate on the use of the bonus (Gagliarducci and Guiso, 2015; Pinotti, 2015), we repeat the above estimation on monthly spending for food eaten at home (Table 2). While the effect of the bonus on total consumption is generally not statistically significant and oscillates in magnitude, when considering food expenditure the effect of the bonus is quite consistent throughout the specifications: in particular, for every additional euro received as bonus, roughly 31.5 cents are spent on food consumed at home. This is in line, although slightly larger, with the results by Neri et al. (2017). In a nutshell, we find evidence that the

Table 2
Effect of the bonus on food consumption.

	(1)	(2)	(3)	(4)	(5)	(6)	(7)
bonus_amount	0.497***	0.207**	0.255***	0.263***	0.265***	0.262***	0.315***
	(0.101)	(0.093)	(0.094)	(0.094)	(0.094)	(0.094)	(0.096)
Income	0.005***	0.004***	0.004***	0.004***	0.004***	0.004***	0.004***
	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)	(0.0003)
ncomp		114.2***	109.5***	111.5***	111.3***	102.3***	102.8***
		(11.19)	(10.62)	(10.72)	(10.74)	(11.38)	(11.07)
ncomp2		−7.850***	−6.868***	−7.040***	−7.007***	−6.072***	−5.951***
		(1.890)	(1.796)	(1.793)	(1.795)	(1.816)	(1.754)
Age			1.936***	2.097***	2.109***	2.082***	2.237***
			(0.259)	(0.286)	(0.285)	(0.284)	(0.285)
Diploma				11.98*	11.94*	12.01*	9.926
				(6.892)	(6.893)	(6.888)	(6.803)
Degree				16.45	16.00	16.25	17.90
				(11.46)	(11.48)	(11.50)	(11.43)
Male					−4.279	−6.773	−1.840
					(5.911)	(6.000)	(5.981)
Married						17.51**	21.95***
						(7.633)	(7.635)
Constant	292.0***	78.00***	−8.416	−24.11	−22.77	−13.87	−30.81
	(9.319)	(13.03)	(15.83)	(18.50)	(18.73)	(19.15)	(25.05)
Regional FE	NO	NO	NO	NO	NO	NO	YES
Observations	3,646	3,646	3,646	3,646	3,646	3,646	3,646
R2	0.269	0.388	0.397	0.398	0.398	0.398	0.416

Notes. Estimation on the SHIW 2014 dataset. Only households with at least one employee. *-**-*** denotes statistical significance at 10%−5%−1%. Robust standard errors in parentheses.

Table 3
Effect of the bonus on consumption: heterogeneity analysis by different definitions of consumption constraints.

	Total non-durable consumption		Food consumption	
	(1)	(2)	(3)	(4)
Panel A	No difficulty	Difficulty	No difficulty	Difficulty
bonus_amount	−0.222	0.503	0.052	0.357***
	(0.492)	(0.333)	(0.161)	(0.117)
Observations	1,357	2,289	1,357	2,289
R2	0.485	0.438	0.385	0.437
Panel B	No difficulty (restricted)	Difficulty (restr.)	No difficulty (restr.)	Difficulty (restr.)
bonus_amount	0.081	0.635	0.151	0.486***
	(0.340)	(0.444)	(0.116)	(0.166)
Observations	2,540	1,106	2,540	1,106
R2	0.478	0.474	0.398	0.450
Panel C	Liquidity unconstrained	Liquidity constrained	Liquidity unconstrained	Liquidity constrained
bonus_amount	0.292	−0.721	0.294***	1.057***
	(0.273)	(2.040)	(0.099)	(0.393)
Observations	3,504	142	3,504	142
R2	0.522	0.365	0.413	0.613

Notes. Estimation on the SHIW 2014 dataset. Only households with at least one employee. All controls of specification (7) in Tables 1-2 included. In Panel A, Difficulty = 1 if the household reports making ends meet with great difficulty, with difficulty or with some difficulty. In Panel B, Difficulty = 1 if the household reports making ends meet with great difficulty or with difficulty. In Panel C, Liquidity constrained = 1 if the household was at least partially rejected a request for a mortgage, or would have liked to apply for it but had not because of the assumption to be rejected. *-**-*** denotes statistical significance at 10%−5%−1%. Robust standard errors in parentheses.

bonus has an effect on food consumption. The effect is statistically significant even when we introduce additional controls and its economic magnitude is stable across specifications. In contrast, the effect of the tax rebate on total consumption is not detectable in our data. In the remainder of the paper, we will mainly focus on food consumption in documenting the current misallocation of the bonus.

Using the specification in column (7) of the above tables, we then investigate whether the bonus has a heterogeneous effect on consumption according to the different proxies for being consumption constrained (Table 3). In Panel A and B we focus on the self-reported difficulties to make ends meet. In Panel A, difficulty is equal to 1 if the household makes ends meet with great difficulty, with difficulty or with some difficulty. Noting that a considerable number of households reported to face some difficulty getting through the month, we slightly modify the definition of difficulty in Panel B, so to include only households making ends meet with great difficulty or with difficulty. In both specifications, total consumption does not seem to significantly react to the tax rebate. Nevertheless, the bonus seems to increase food consumption for all households and the effect is larger and statistically significant for families facing difficulties. In Panel C, we estimate the effect of the bonus among households that report to be liquidity constrained and households that do not. Differently from the difficulty in making ends meet, this indicator provides an assessment of the overall households' wealth, rather than income only. The bonus has a significant and positive effect only on food consumption for both constrained and unconstrained households, and such increase is larger for constrained families. In our data, being consumption constrained is a relevant phenomenon, characterizing about 60% of households; on the other hand, only 4% of households report themselves to be liquidity constrained.

In the remainder of the paper, our preferred indicator of consumption constrained households is the difficulty to make ends meet. First of all, this indicator comes from a questionnaire variable available for each wave in the same manner. This allows to predict the same variable in both the 2010–12 sample and the 2014 one, thus making the ML targeting exercise (see Section 5) feasible.[10] Secondly, we prefer to work with a reasonable number of observations of consumption constrained households. This excludes the liquidity constrained indicator that identifies few households as such, as shown in Panel C of Table 3. [11] Thirdly, we wish to use a variable that allows us to estimate an effect of the bonus on consumption which is in line with the other papers focused on estimating the effect of the bonus. As one can see from Table 3, the results in Panel A obtained using our preferred indicator are the closest to the ones obtained by Neri et al. (2017). Fourthly, we prefer to use the unrestricted version of the difficulty to make ends meet indicator, although we discuss what happens to our results when more restrictive proxies are instead used.

5. Targeting analysis

In what follows, we illustrate the procedure used to select the variables to be employed in the prediction exercise (Section 5.1); we next describe the features of the prediction exercise itself (Section 5.2); we finally show and comment the empirical findings (Section 5.3).

5.1. Variable selection

For the targeting analysis, we want to identify in the 2014 sample those households that are more likely to be consumption constrained. In this way, we can evaluate the effectiveness of the current allocation and suggest some possible alternatives. Although we observe consumption constrained households also in the 2014 sample, this information is not useful for targeting: we need to rely on information available prior to the start of the policy, which could have been used by the policy maker. We therefore focus on a pooled dataset of the 2010 and 2012 waves.[12] We only maintain one feature of the current policy, that is the focus on employees. Hence, we select only households with at least one employee. On this sample, we estimate models that allow us to predict the consumption constrained status on the basis of a set of observable characteristics. We consider a set of variables that are recorded in both the pooled 2010–12 dataset and the 2014 one, so as to predict the consumption constrained status in 2014 using the prediction model estimated on the 2010–12 dataset.[13] We dismiss all the variables that are excluded for collinearity reasons by running a simple regression of our proxy for the consumption constrained status on the covariates set. The complete list of variables used for prediction can be found in Supplementary Table A.2. They essentially refer to household income, wealth, and demographic characteristics.

5.2. Prediction

ML techniques rely on highly flexible functional forms. Nevertheless, allowing for more flexibility in the model, improves the in-sample fit at the cost of reducing the out-of-sample fit (over-fitting). To account for this trade-off, each ML algorithm comes with a regularization of the complexity level, which is chosen by an empirical tuning, usually based on 10-fold cross-validation (see Hastie et al., 2009). Given that the main purpose is out-of-sample prediction, we estimate and tune the models on a training subsample, composed of a randomly selected 2/3 sample of the 2010–12 pooled dataset. The remaining 1/3 of such dataset constitutes the testing subsample.

Our main ML algorithm is the decision tree (Hastie et al., 2009). Decision trees are particularly appropriate for applications in which the assignment mechanism needs to be transparent; for instance, when the results need to be shared in order to facilitate decision making (Lantz, 2013). As it will be clear in Section 5.3, the output of a decision tree algorithm can be easily described. The algorithm divides data into progressively smaller subsets to identify patterns that can be used for predicting a specific binary output. In our case, the algorithm creates a decision rule which partitions the observations

[10] SHIW 2010 and 2012 contain hypothetical questions about the propensity to consume out of an income shock that are possibly closer to our scope and that have been used also by Jappelli and Pistaferri (2014). However, the question changes significantly between the two waves and it is not available in 2014.

[11] Nevertheless, about 90% of the liquidity constrained households also report to have difficulties in making ends meet.

[12] As mentioned before, the sample includes a longitudinal component, but we ignore it because the sample size would be too small for a reliable targeting analysis. Therefore, i univocally identifies a household-year pair.

[13] Athey (2018) highlights that ML predictions might have poor stability and robustness properties in response to variations in sample or variations in the environment. In our case, the relationship between selected predictors and predicted status appears to be quite stable (see footnote 18).

according to the consumption constrained status on the basis of the values of the vector of observable covariates (z_i). Non linearities and interactions are captured by the sequence of splits. Following a top-down approach, at each step the algorithm selects a variable z_{gi} from z_i and splits the observations into two groups according to a threshold $\bar{z}_g$ (or according to a subset of values in case of a multinomial discrete variable). Both the variable used to split and the threshold are chosen to obtain the largest possible reduction in heterogeneity (impurity) of the variable to be predicted (Siroky, 2009).[14] The algorithm then proceeds to the next step by further splitting the sub-samples at each terminal node. It stops when the degree of impurity of a terminal node is as low as possible. A high number of levels in a tree is likely to overfit the data. This could lead to a model which performs very well in the training sample, but gives highly imprecise predictions out-of-sample (Athey and Imbens, 2017; Lantz, 2013; Breiman et al., 1984). A solution to this problem is to reduce the complexity of the tree by setting a complexity parameter (cp) and use it to prune the tree. We choose the optimal cp by using a rule of thumb suggested in the literature and based on cross-validation (Hastie et al., 2009).[15]

We compare the findings obtained with the decision tree with those deriving from other ML algorithms, the k-Nearest Neighbors (k-NN) and random forest, as well as a standard linear probability model (LPM).[16] In the k-NN algorithm (Lantz, 2013), the trade-off between overfitting and out-of-sample prediction is solved by choosing the optimal number of neighbors (i.e. the level of k). For each observation in the testing sample, the algorithm identifies the k closest observations from the training sample (the so-called nearest neighbors) and assigns a prediction on the basis of a majority rule, i.e. it takes as prediction the most frequent outcome among those of the nearest neighbors. We chose the optimal number k of neighbors by using the 10-fold cross-validation. The random forest algorithm explores a richer set of possible models. Essentially, it estimates a large number of trees on a series of new samples generated by randomly drawing (with replication) from the original sample (i.e. bootstrapping), using only a randomly selected subset of the regressors for each tree. To obtain the final prediction for each observation, random forest applies a majority vote rule across the predictions generated by each tree. Intuitively, the algorithm works as a decision tree that moves around over lots of regressors. Although the random forest performance improves with strong and moderately important predictors, the algorithm is not free from the risk of averaging over noise as it may also select regressors that are highly correlated with predictors. Therefore, one should use random forest only if the number of regressors is really big, which is not our case. Following the work of Chandler et al. (2011), who estimated a model to predict highest risk youth for anti-violence interventions, we also make use of a LPM prediction. To make it more comparable with ML predictions, we include all the variables depicted in Supplementary Table A.2, the squares and cubes of the continuous variables, plus all the interactions between themselves and all the interactions between the continuous and discrete covariates. In the case of LPM, the prediction is continuous, so we consider the dummy consumption constrained having the value of 1 if the predicted probability is larger than 0.5. k-NN, random forest and LPM are used essentially to probe the prediction quality of the decision tree. In our case, they cannot be considered as real alternatives, as we are looking for a transparent assignment mechanism.

5.3. Empirical findings

The decision tree leads to the assignment mechanism shown in Fig. 6. It depends on few variables, essentially referring to household income and wealth. The targeted households would be: (a) those that have financial assets lower than 13,255 euro; among these ones, the consumption constrained households are those that either perceive income lower than 36,040 euro yearly or those that earn more than 36,040 euro but such that the maximum income perceived within the household is lower than 34,500 euro; (b) those that have financial assets higher than 13,255 euro; among these ones, the consumption constrained households are those that earn less than 52,591 euro yearly and have an income from financial assets lower than 432.9 euro together with a minimum income perceived within the household lower than 13,895 euro.

Next, we use k-NN, random forest, and LPM. These methods do not provide clear insights on which characteristics of the household are pivotal in the selection rule. Table 4 compares the performance of the models in terms of correctly predicting the consumption constrained status. Notwithstanding its simplicity, the decision tree correctly identifies 74% of the observations, a percentage very close to that of its alternatives (73%, 77% and 75% for the k-NN, the random forest and the LPM, respectively).[17]

Using the 2014 data, we proceed to estimate the effect of the 80 euro rebate distinguishing between the households that, according to our decision tree assignment, should have received the bonus being predicted as consumption constrained, and those that should have not.[18] Table 5 reports the results of the estimation. The effect of the bonus on food consumption is

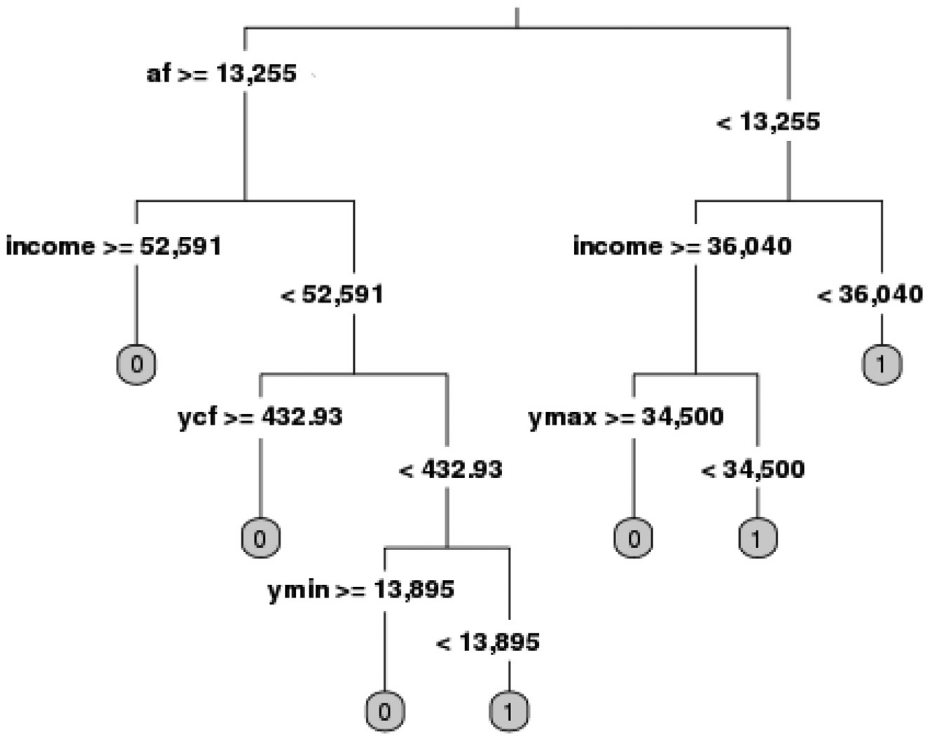
[14] We use the R package "rpart" written by Therneau et al. (2015). The degree of impurity at each node (leaves) is measured using a heterogeneity index.

[15] First, the complexity parameter associated to the smallest 10-fold cross-validation error (say $errmin$) is found. Then, the optimal cp is the one that has a cross-validated error which is the closest to $errmin + standard\_error(errmin)$. The rule of thumb leads to a simpler tree because the cross-validation error curve tends to be flat around its minimum, hence there is a small gain in picking exactly the minimum, while there is a higher risk of over-fitting.

[16] We use the R packages "knn" (Venables and Ripley, 2002) and "randomForest" (Liaw and Wiener, 2002).

[17] Another strategy could be to use the decision tree for variable selection and then the LPM as prediction model. In such a case, the out-of-sample misclassification rate is 28.3% which is higher than 25.9% obtained with the decision tree. As for comparison, the LPM specification estimated on the full set of variables and interaction terms has a miss-classification rate equal to 25.0%.

[18] Since we are using 2010–12 information to predict 2014 consumption constrained households, we also investigate whether the association between the actual consumption constrained status and tree-selected predictors is stable. We run two separate LPM regressions for the 2010 and the 2012 sub-samples, using the dummy for difficulty in making ends meet as dependent variable and the variables selected by the tree as covariates. The relationship



Legend	
af	household yearly financial assets, euro
income	household yearly disposable income, euro
ycf	household income from financial assets, euro
ymin	minimum individual labor income within the household, euro
ymax	maximum individual labor income within the household, euro

Fig. 6. Decision tree output.

Table 4
Decision tree, k-NN, LPM and random forest models performance compared.

		Real status Not consumption constrained	Consumption constrained	Total
Panel A: decision tree				
	Not consumption constrained	608	232	840
Predicted status	Consumption constrained	447	1,334	1,781
	Total	1,055	1,566	2,621
	% Correctly predicted	57.6%	85.2%	74.1%
Panel B: k-NN				
	Not consumption constrained	593	244	837
Predicted status	Consumption constrained	462	1,322	1,784
	Total	1,055	1,566	2,621
	% Correctly predicted	56.2%	84.4%	73.1%
Panel C: random forest				
	Not consumption constrained	680	218	898
Predicted status	Consumption constrained	375	1,348	1,728
	Total	1,055	1,566	2,621
	% Correctly predicted	64.5%	86.1%	77.4%
Panel D: LPM				
	Not consumption constrained	608	208	816
Predicted status	Consumption constrained	447	1,358	1,805
	Total	1,055	1,566	2,621
	% Correctly predicted	57.6%	86.7%	75.0%

Notes. Out-of-sample estimation on the testing subsample of the 2010–12 pooled SHIW dataset (only households with at least one employee). Consumption constrained = 1 if the household makes ends meet with great difficulty, with difficulty or with some difficulty.

Table 5
Effect of the bonus on consumption by predicted consumption constrained status (decision tree).

	Total consumption (1)	(2)	Food consumption (3)	(4)
Predicted status:	Not consumption constrained	Consumption constrained	Not consumption constrained	Consumption constrained
bonus_amount	−0.527	0.710**	0.009	0.369***
	(0.563)	(0.315)	(0.184)	(0.111)
Observations	1,146	2,500	1,146	2,500
R2	0.459	0.415	0.356	0.442

Notes. Estimation on the SHIW 2014 dataset. Only households with at least one employee. All controls of specification (7) in Tables 1-2 included. Consumption constrained = 1 if according to the decision tree algorithm the household is predicted to make ends meet with great difficulty, with difficulty or with some difficulty. *-**-*** denotes statistical significance at 10%–5%–1%. Robust standard errors in parentheses.

Table 6
Decision tree rule: predicted consumption constrained status and bonus recipients.

		Predicted status Not consumption constrained	Consumption constrained	Total	% Overlapping
Bonus recipient	Not recipient	715	1,446	2,161	33.0%
	Recipient	431	1,054	1,485	71.0%
	Total	1,146	2,500	3,646	
	% Overlapping	62.4%	42.2%		48.5%

Notes. Estimation on the SHIW 2014 dataset. Only households with at least one employee. Consumption constrained = 1 if the household is predicted to make ends meet with great difficulty, with difficulty or with some difficulty.

positive and significant for the households that would have been targeted with our assignment rule. The effect is, instead, neither statistically nor economically different from zero for households that received the bonus without being consumption constrained according to the decision tree rule. In particular, households predicted to be consumption constrained spend on average 36.9% of the bonus in food consumption. This share is very close to the one estimated by splitting the sample using the actual reported status in the 2014 data (see Table 3).[19]

Table 6 provides the percentage of overlap between predicted status (i.e. being consumption constrained or not) and the receipt of the bonus. The overlap includes households that: (i) both receive the bonus and are predicted to be consumption constrained, and (ii) both do not receive the bonus and are predicted to be not consumption constrained. This fraction is around 49%. Given that we find evidence of an impact on consumption only among those predicted to be consumption constrained, this implies that there were margins to improve the effectiveness of the program.

In order to capture the misallocation, we focus on a measure of spending inefficiency due to the actual allocation rule. As shown in Table 6, 71.0% of the households that receive the bonus are predicted to be consumption constrained by the decision tree algorithm. Our spending inefficiency measure refers to the remaining 29.0%. We look at the amount that was spent for the bonus recipients that the decision tree does not identify as consumption constrained households. We compute such a measure as follows. Let A be the set of bonus recipients in our dataset, and B the subset of A made up of predicted consumption constrained households. The total expenditure for the tax rebate is given by

$$E_{total} = \sum_{i \in A} bonus\_amount_i \tag{3}$$

while the "correct" expenditure (namely, the amount spent for the predicted consumption constrained households) is given by

$$E_{correct} = \sum_{i \in B} bonus\_amount_i \tag{4}$$

Therefore, the percentage of expenditure that has been misallocated can be computed as

$$\frac{E_{total} - E_{correct}}{E_{total}} \tag{5}$$

This share turns out to be equal to 29.5% of the total expenditure. If households that are not predicted to be consumption constrained were excluded from the bonus, 29.5% of the 7 billion transfer entailed under the current policy (roughly 2

between the observables and the consumption constrained status appears to be quite stable, as coefficients change only marginally. Results are presented in Supplementary Table A.3.

[19] One issue is that predictors and the consumption constrained status are both measured at the same time. In principle, one would predict the status with variables that have been already observed at the time the policy is implemented. Our data do not allow us to follow such a strategy. However, note that the selected predictors such as income and wealth are characterized by a high degree of persistence. In particular, we use the panel component of the dataset and regress each predictor measured in 2014 on its 2010–12 average value. Such an estimate is roughly 0.9 for the two main predictors (income and financial assets).

billion euro) could be saved without changing the impact on consumption. Alternatively, in order to increase the impact on consumption, this amount could be reallocated to those households that are predicted to be consumption constrained but did not receive the bonus. One possibility is to endow this group with a transfer which is set to be equal to the per capita transfer received by households belonging to B (i.e. roughly 57 euro). In such a case, keeping fixed the total public expenditure, we could reach 31% of predicted consumption constrained households that did not receive the bonus (60% of the households we predict as consumption constrained would be endowed with a bonus).[20] As shown in Table 5, for predicted consumption constrained households, the estimated coefficient of the bonus on food consumption is 0.369. Assuming that the behavior of predicted consumption constrained households who are not bonus recipients under the actual rule would be the same as the behavior of those who receive it, reallocating the saved resources to predicted consumption constrained households improves the bonus impact on food consumption by about 760 million euro (+41.8% with respect to the current impact).

In short, the current scheme could be improved: (i) by excluding those that are not predicted as consumption constrained to get the same impact on consumption as the current rule but saving resources; (ii) by reallocating the misallocated expenditure to predicted consumption constrained households that did not receive the bonus to obtain a greater impact on consumption.

6. Definitions, extensions and concerns in using ML for policy choices

In what follows, we first discuss the use of alternative definitions of consumption constrained households and whether we could further tune the model based on our preferred definition (Section 6.1); we then illustrate the possible extension of our exercise to retirees (Section 6.2); next, we discuss the omitted payoffs of our main targeting exercise (Section 6.3); finally, we address the issues related to data requirements, transparency and manipulation (Section 6.4).

6.1. The consumption constrained group: definitions and further tuning

The group of predicted constrained households is quite large. In the 2014 wave, the number of households that are predicted to be in such a status is 2,500, which is around two thirds of the entire sample. As we are evaluating how to allocate scarce resources, a policy maker may want to focus on a smaller group.

One way to do so could be focusing on more restrictive definitions of consumption constrained households (see Section 4). Using them, the percentage of ML targeted households is obviously lower, thus increasing the percentage of expenditure that ends up to be misallocated. In particular, the SHIW question on making ends meet allows us to work with two alternatives. We can consider as consumption constrained households: (i) those who make ends meet “with difficulty” or “with great difficulty” (therefore excluding those with only “some difficulties”); (ii) only those who report to make ends meet “with great difficulty”. As for the first alternative definition, applying the decision tree rule to the 2014 sample leads to 673 households predicted to be consumption constrained, against 2,973 not consumption constrained; as for the second definition, we end up with 345 predicted consumption constrained households, against 3,301 not consumption constrained. One problem is that there does not seem to be a clear pattern for the impact of the bonus across variously defined groups of consumption constrained households. For instance, the bonus effect on food consumption is actually lower than our baseline (Table 5) for the group “with difficulty” or “with great difficulty” (0.227; s.e. 0.203), but it is higher for the group that includes only households “with great difficulty” (0.491; s.e. 0.267). Moreover, when we use these alternative definitions, we find evidence of a positive and statistically significant effect of the bonus on food consumption also among those that are predicted as not constrained (while the effect is zero in the baseline results of Table 5). Taken together, these two pieces of evidence suggest that our preferred definition, although less restrictive than the others, is the only one that allows to put together all those households that, when targeted under the policy, react to the bonus with a positive and significant consumption response. The choice to target a smaller group of people, as those identified by the restrictive definitions, is therefore a sheer political matter with respect to which we remain neutral.[21]

Even if one agrees that our consumption constrained definition should be preferred, he or she may still think that the number of predicted consumption constrained households is too large (2,500 with respect to 2,289 actually declaring to be consumption constrained in 2014). In order to reduce this number, we can tune the algorithm by increasing the penalty for not consumption constrained households that are predicted to be constrained (i.e. the false positives). We chose to set such a penalty to 1.4 and a penalty equal to 1.0 for consumption constrained households that are predicted to be not constrained (i.e. the false negatives). By doing so, the fraction of false positives in the 2010–12 testing sample falls to roughly 13% (it is about 17% with the baseline decision tree). Looking at the 2014 sample, the chosen penalties are such that the number

[20] A different strategy, that would still be consistent with a binding resource constraint, could be to give the (same amount of) bonus to all the predicted consumption constrained households. In order to reach this goal, the per capita amount of the bonus should be reduced to roughly 34 euro for those households that are already included in the policy and are consumption constrained. However, as the characteristics of the current transfer may satisfy other policy goals over which we remain agnostic, we prefer to focus on a proposal that leaves the amount of the bonus unchanged for the households that are already included in the actual policy.

[21] In the case of the alternative definitions, financial assets and disposable income remain the most important predictors of the consumption constrained status according to the decision tree algorithm.

Table 7
Retirees: effect of an income shock on food consumption by potential recipients.

	Food consumption	
	(1)	(2)
Panel A: allocation rule based on income quartiles	Not recipient	Recipient
Income	0.00302***	0.00325***
	(0.00047)	(0.00074)
Observations	2,380	2,373
R2	0.427	0.427
Panel B: allocation rule based on the consumption constrained status	Not recipient	Recipient
Income	0.00197***	0.00619***
	(0.00040)	(0.00044)
Observations	1,778	2,975
R2	0.323	0.474

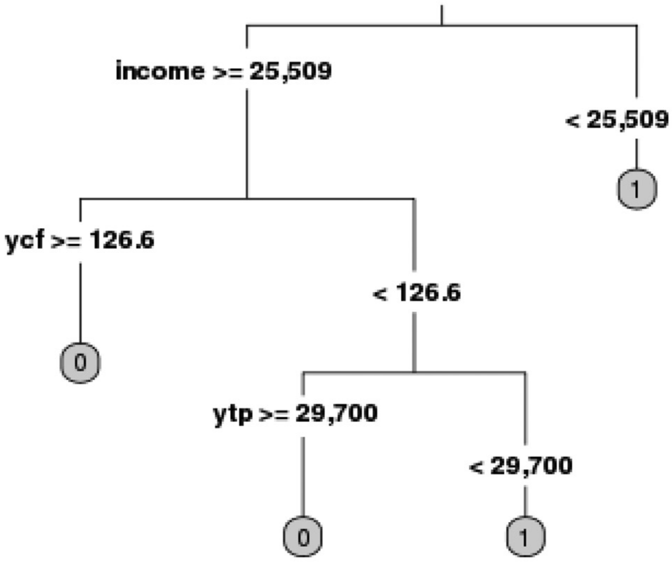
Notes. Estimation on the SHIW 2014 dataset, selecting only households with at least one retiree. All controls of specification (7) in Tables 1-2 included, plus a dummy for households already receiving the current bonus. In Panel A, Recipient = 1 if average income from pension within the household is between the second and third quartile of the distribution. In Panel B, Recipient = 1 if the household makes ends meet with great difficulty, with difficulty or with some difficulty. *-**-*** denotes statistical significance at 10%–5%–1%. Robust standard errors in parentheses.

of predicted consumption constrained households (2,221) ends up to be roughly in line with the number of households actually reporting to be so. Using this new group of predicted constrained households, the estimated bonus effect on food consumption is equal to 0.400 (statistically significant) which is higher, although only slightly, than the one we estimate in our baseline targeting analysis (0.369). The bonus effect on food consumption for households predicted to be not constrained remains close to zero. Finally, as could be expected, the fraction of misallocated expenditure increases to 38%. All in all, this suggests that tuning the algorithm to reduce the rate of false positives could, in this case, further increase the gains from ML targeting. However, the fraction of false negatives increases to 12% (it is roughly 9% in the baseline decision tree) meaning that more consumption constrained households would be predicted to be not constrained and, therefore, erroneously excluded from the policy scenario we suggest in Section 5.3. Taking a position on the weight to be given to false positives vs false negatives should be underpinned by a specific welfare function chosen by the policy maker. For this reason, in our baseline analysis, we prefer to give equal weight to false positive and false negative cases. Nevertheless, the above exercise is useful to show that our method is flexible enough to accommodate different policy objectives and that we are at most understating the potential gains of ML targeting.

6.2. Retirees

We now sketch how a prediction exercise may be implemented for the retirees, which have not been considered under the scheme so far but may be thought as a possible additional group to be targeted under the policy. We consider the households in which there is at least one retiree and create an allocation rule which mostly resembles the one currently applied to employees, given that also for the pensioners we do not observe gross income. To this aim, we consider the distribution of the average income from pensions within the household and we divide it into quartiles. In line with the actual bonus allocation rule, we then exclude the poorest and richest households. In other words, we assume that the bonus recipients are those households whose average income from pension is between the second and the third quartile of the distribution. The income thresholds thus obtained are consistent with those defined for the actual bonus.

In order to check whether there is any baseline difference in food consumption behavior between potential recipients and not recipients, we regress the food expenditure on income and the same set of controls we used in our main exercise. Table 7 (Panel A) shows that such difference seems not to exist.[22] We then assume that potential recipients are those households who are consumption constrained according to our baseline definition (i.e., households that make ends meet with some difficulty, with difficulty or with great difficulty). Panel B provides evidence of a different food consumption behavior between consumption constrained and not consumption constrained households. Namely, the effect of an income perturbation on food expenditure is higher for consumption constrained households than for the ones that are not. We then apply the decision tree algorithm to the retirees sample. Fig. 7 shows the profile of the predicted consumption constrained retirees. The targeted households, in this case, are those perceiving income lower than 25,509 euro yearly or those that earn more than 25,509 euro but have income from financial assets lower than 126.6 euro together with an household yearly income from retirement lower than 29,700 euro.



Legend	
income	household yearly disposable income, euro
ycf	household income from financial assets, euro
ytp	household yearly income from retirement, euro

Fig. 7. Retirees: decision tree output.

Table 8
Retirees: effect of an income shock on consumption by predicted consumption constrained status.

	Total consumption (1)	(2)	Food consumption (3)	(4)
Predicted status	Not consumption constrained	Consumption constrained	Not consumption constrained	Consumption constrained
Income	0.01147***	0.01851***	0.00155***	0.00544***
	(0.00180)	(0.00121)	(0.00037)	(0.00048)
Observations	1,475	3,278	1,475	3,278
R2	0.353	0.463	0.235	0.453

Notes. Estimation on the SHIW 2014 dataset, selecting only households with at least one retiree. All controls of specification (7) in Tables 1-2 included, plus a dummy for households already receiving the current bonus. Consumption constrained = 1 if the household is predicted to make ends meet with great difficulty, with difficulty or with some difficulty. *-**-*** denotes statistical significance at 10%–5%–1%. Robust standard errors in parentheses.

Table 8 reports the results of the estimation using the predicted consumption constrained status. Again, the effect of an income increase on food consumption is higher for those households that are predicted to be consumption constrained. It is worth noting that targeting these households using the decision tree algorithm delivers a positive and statistically significant effect of an income increase on total consumption as well.

6.3. Omitted payoffs

Our decision tree rule is designed to increase the consumption impact of the bonus without considering other possible outcomes. However, as discussed by Kleinberg et al. (2018), any targeting rule might have "omitted payoffs" because the new allocation may indirectly prioritize households with specific characteristics along other correlated dimensions. One relevant issue is whether a targeting rule based on the ML predicted status would have a different redistributive effect. With respect to the current allocation, our proposal is more tailored towards households with relatively lower income (among those with

[22] We also consider the possibility to allocate the bonus to the poorest retirees and thus re-run the above exercise assuming that recipients are those households whose average income from pension is in the first quartile of the distribution. Even in this case, we find no evidence of a statistically significant difference in food consumption behavior between potential recipients and not recipients. This reinforces our argument that allocating the bonus on the basis of income only may be misleading.

Table 9
Income by predicted consumption constrained status and bonus receipt.

Bonus recipient	Predicted status	Average household income	Average equivalent household income	Observations
Recipient	Not consumption constrained	47,884	24,818	431
Recipient	Consumption constrained	29,149	15,586	1,054
Not recipient	Not consumption constrained	58,817	31,844	715
Not recipient	Consumption constrained	28,583	15,874	1,446

Notes. Estimation on the SHIW 2014 dataset. Only households with at least one employee. Consumption constrained = 1 if the household is predicted to make ends meet with great difficulty, with difficulty or with some difficulty. Equivalent income is the ratio between total household income and the number of equivalent adults. The latter is calculated using the OECD-modified equivalence scale which assigns a value of 1 to the head of household, of 0.5 to each additional member over the age of 14 and of 0.3 to those below that age. The households income is calculated net of the bonus.

at least one employee). It therefore entails a redistributive effect which is likely to be stronger than the current one. To document this point, in Table 9 we split the 2014 sample by actual bonus receipt and ML predicted status. Within those that actually receive the bonus, the households that are predicted to be not consumption constrained have larger income. This is true also if we consider income in equivalent terms (by dividing it by the modified OECD scale, as standard in the inequality literature). Similar differences apply to the ML vs non ML targeted households among those that did not receive the bonus in 2014. If we re-designed the allocation rule by using ML predictions, as suggested in Section 5.3, we would prioritize households that have relatively lower income.

It is more complicate to understand how the ML targeting would change the impact of the bonus on labor supply. The actual effect of the current policy on employment is, from a theoretical standpoint, ambiguous. On the one hand, the bonus reduces the overall taxation on wages for eligible individuals, hence it increases the probability that an individual accepts an employment offer with a given salary. On the other hand, the bonus may have negative effects on the intensive margin because, for individuals whose gross wage is between 24,000 and 26,000, the effective tax rate, which accounts for the decrease in the bonus amount, is extremely high. Thus, the overall effect on labor supply of the current scheme depends on the combination of the two mechanisms. Also the rule that we suggest influences labor supply through these two channels. In the predicted consumption constrained group, which includes individuals with relatively lower income and wages, the increase in the probability of accepting an employment offer is likely to be larger than that associated to the actual policy. However, within this group, the contraction in the intensive margin might be stronger. Furthermore, looking at the household level response, our proposed rule has a drawback with respect to the current one. Being designed on family characteristics, eligibility depends not only on the individuals' income, but on the household's. As a result, it may have a negative effect on the second earner, usually the female, as her decision to work may lead the household not to be eligible anymore. A full discussion requires a behavioral micro-simulation model that simulates the agents' behavioral response accounting for the heterogeneity in their budget constraints (see, for instance, Pacifico, 2009). In any case, by simply excluding those that are predicted to be not consumption constrained (Section 5.3), we would still obtain significant fiscal savings as 29.5% of the allocated expenditures did not serve to increase consumption. This public budget could be devoted to other targets, such as cutting taxes on both households and firms.

6.4. Data requirement, transparency and manipulation

The decision tree rule is based on information at household level that, at least in principle, is observable by a policy maker. As a matter of example, the equivalent economic situation indicator (i.e., the so called "ISEE") enables the policy makers to collect information on income and wealth at household level. We are aware that implementing the targeting rule we suggest may increase the costs of the policy in the short term because it would require, using the same example, to know the ISEE of all Italian households. However, the use of household-level information is also in line with other recent proposals to review some assistance benefits policies aiming at the use of eligibility criteria that approximate the ISEE or, more generally, the household economic condition. In short, data defined at the household level are going to be collected anyway to comply with a more efficient welfare system. Furthermore, the decision tree rule is based on few variables and therefore requires less information than the rules suggested by the other methods that we explored.

Note also that having only a subset of the (few) variables included in the decision tree assignment rule will deliver lower but still sizable benefits. Indeed, a useful feature of the decision tree algorithm is the possibility to compute the fraction of households that would be incorrectly identified as consumption constrained by observing only a subset of characteristics among those involved in the tree. For instance, let us assume that the policy maker can observe household financial assets and disposable income only. In this case, her decision rule to identify consumption constrained households could be based only on the financial assets and income thresholds given by the tree. In terms of Fig. 6, consumption constrained households would be those that have financial assets lower than 13,255 euro and disposable income lower than 36,040 euro; not consumption constrained households would be those that have financial assets at least equal to 13,255 euro and disposable income at least equal to 36,040 euro yearly. Clearly, these groups only partially overlap with the groups of predicted consumption constrained and predicted not consumption constrained households identified through the use of all the variables involved in the tree. Using a decision rule based on financial assets and disposable income only, 21% of the households in the 2014 sample would be allocated to a status that does not correspond to the one predicted by the use of all the variables. If the policy maker observes the maximum income perceived within the household too, and constructs a decision rule also based on this variable using the thresholds given by the tree, then the fraction of incorrectly allocated households decreases to 5%. Finally, the fraction of incorrectly identified consumption constrained households is obviously zero in case both income from financial assets and minimum income perceived within the household are observable, because the decision rule now coincides with the entire tree.

When targeting a policy intervention, the policy makers might have to balance various considerations. A very relevant issue in social policy is the transparency of the targeting rule. As any allocation mechanism ends up in people that are excluded, it is crucial to focus on a transparent assignment mechanism (Athey, 2017, 2018; OECD, 2017). In this respect, the decision tree algorithm allows to summarize the assignment mechanism in a simple figure, as the one of Fig. 6. Alternative methods are more opaque. As for the linear probability model, in order to allow for a flexible data interpolation, we need to impose a deep structure on the data through a large set of interactions that have questionable (or, even, absent) economic meaning as well as no clear interpretation for a stakeholder. As it happens in our case (Table 4), policy targeting could be improved by relying on black box ML methods such as random forest, which usually provides better predictions. Therefore, choosing transparency may lead to inefficiencies which should be "priced". In our context, the "price of transparency" is not so large. Indeed, if we drop the requirement of transparency and rely on a black box ML algorithm such as random forest, the futile expenditure associated to its predictions is 31.7%, slightly higher than the percentage obtained with the decision tree.[23]

As a last point, one may argue that people may react by manipulating the subset of endogenous variables selected by the tree when they are used by the policy maker to distribute the bonus. We believe manipulation is very unlikely in our context. In case of decision tree targeting, the potential recipient should manipulate more than one single variable (say, both financial assets and income, in the simplest case) and the manipulation should involve more than one single family member (given that we are considering mainly household level variables and that both the minimum and maximum income are, taken together with all the other variables, good predictors of the consumption constrained status). Manipulating one single variable, as the case of the actual measure, is instead, more likely, and even more realistic if it refers to individual features.

7. Conclusions

ML provides economists with a powerful toolbox to deal with predictive tasks. We investigate how off-the-shelf ML methods can help improving the design of a public scheme in order to increase its effectiveness. We focus on a massive tax rebate program implemented in Italy in 2014, having as its main objective the households consumption spur, and make use of ML prediction techniques to identify the households that would have benefited more from the program in terms of increased consumption. The gain from such ML based targeting is large. If the recipients of the bonus had been determined using ML, the overall food consumption would have been 41.8% larger (about 760 million euro) with respect to what happens with the actual rule. Alternately, 29.5% of the funds earmarked to the policy (about 2 billion euro) could have been saved without reducing the overall consumption expenditure, as they were channeled to households not targeted by ML. We are aware that ML algorithms should not be used lightly. A relevant issue is the transparency and the interpretability of the targeting rule. In our case, we seek transparency by choosing the decision tree as favorite ML algorithm. All in all, the loss in predictive performance with respect to more complex ML methods is negligible. Another question concerns the possibility that the ML targeting rule might end up having other effects than those explicitly addressed, which might be either desirable or not. Along these lines, we discuss whether the targeting rule based on ML would have different redistributive and labor supply effects with respect to the current allocation.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.jebo.2018.09.010.

References

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.
Athey, S., Imbens, G., 2017. The state of applied econometrics: causality and policy evaluation. J. Econ. Perspect. 31 (2), 3–32.
Baldini, M., Giarda, E., Olivieri, A., Pellegrino, S., Zanardi, A., 2015. Il "bonus" degli 80 euro: caratteristiche ed effetti redistributivi. Rivista di Diritto Finanziario e Scienza delle Finanze 74 (1), 3–22.
Breiman, L., Friedman, J., Stone, C.J., Olshen, R.A., 1984. Classification and Regression Trees. Wadsworth, Belmont.
Chalfin, A., Danieli, O., Hillis, A., Jelveh, Z., Luca, M., Ludwig, J., Mullainathan, S., 2016. Productivity and selection of human capital with machine learning. Am. Econ. Rev. 106 (5), 124–127.
Chandler, D., Levitt, S.D., List, J.A., 2011. Predicting and preventing shootings among at-risk youth. Am. Econ. Rev. 101 (3), 288–292.

[23] One may also think to estimate a simple LPM based on the few variables selected by the decision tree. We have already argued that such a strategy has a poor out-of-sample forecasting performance (see footnote 17). Additionally, it also ends up in a too low price for transparency. Indeed, the futile expenditure associated to this strategy is 16.1%, a much lower value than the one obtained with our preferred decision tree.

Gagliarducci, S. and Guiso, L. (2015). Gli 80 euro? Spesi al supermercato. Lavoce.info. Available at: http://www.lavoce.info/archives/36685/gli-80-euro-spesi-al-supermercato.

Hastie, T., Tibshirani, R., Friedman, J., 2009. The Elements of Statistical Learning - Data Mining, Inference, and Prediction. Springer, New York.

Jappelli, T., Pistaferri, L., 2014. Fiscal policy and MPC heterogeneity. Am. Econ. J.: Macroecon. 6 (4), 107–136.

Kang, J.S., Kuznetsova, P., Luca, M., Choi, Y., 2013. Where not to eat? Improving public policy by predicting hygiene inspections using online reviews. In: Proceedings of the 2013 Conference on Empirical Methods in Natural Language Processing, pp. 1443–1448. Available at: https://aclanthology.info/pdf/D/D13/D13-1150.pdf.

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J., Mullainathan, S., 2018. Human decisions and machine predictions. Q. J. Econ. 133 (1), 237–293.

Kleinberg, J., Ludwig, J., Mullainathan, S., Obermeyer, Z., 2015. Prediction policy problems. Am. Econ. Rev. 105 (5), 491–495.

Lantz, B., 2013. Machine Learning with R. Packt Publishing Ltd, Birmingham.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2 (3), 18–22.

McBride, L., 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.

Ministry of Economics and Finance, 2015. Statistiche sulle dichiarazioni fiscali. Analisi dei dati Irpef - Anno d'imposta 2014. Available at: http://www1.finanze.gov.it/finanze2/analisi_stat/v_4_0_0/contenuti/analisi_dati_2014_irpef.pdf?d.

Mullainathan, S., Spiess, J., 2017. Machine learning: an applied econometric approach. J. Econ. Perspect. 31 (2), 87–106.

Neri, A., Rondinelli, C., and Scoccianti, F. (2017). Household Spending Out of a Tax Rebate: Italian "€80 Tax Bonus". Bank of Italy Occasional Paper No. 379, June 2017, Bank of Italy.

Organization for Economic Cooperation and Development, 2017. Algorithms and collusion: Competition policy in the digital age. Available at: www.oecd.org/competition/algorithms-collusion-competition-policy-in-the-digital-age.htm.

Pacifico, D. (2009). A Behavioral Microsimulation Model with Discrete Labour Supply for Italian Couples. CAPP Working Paper No. 65, March 2009, University of Modena and Reggio Emilia.

Pinotti, P. (2015). 80 euro rimasti nel portafoglio. Lavoce.info. Available at: http://www.lavoce.info/archives/37086/80-euro-rimasti-nel-portafoglio.

Rockoff, J.E., Jacob, B.A., Kane, T.J., Staiger, D.O., 2011. Can you recognize an effective teacher when you recruit one? Educ. Finance Policy 6 (1), 43–74.

Siroky, D.S., 2009. Navigating random forests and related advances in algorithmic modeling. Stat. Surv. 3, 147–163.

Therneau, T., Atkinson, B., Ripley, B., 2015. Rpart: recursive partitioning and regression trees - R package version 4.1-10. Available at: https://CRAN.R-project.org/package=rpart.

Varian, H.R., 2014. Big data: new tricks for econometrics. J. Econ. Perspect. 28 (2), 3–28.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S. Springer, New York.