Changing the incentive to pollute: Heterogeneous effects of waste pricing policies

Marica Valente*

Preliminary version, please do not circulate

Abstract This paper estimates heterogeneous causal effects of Pay-As-You-Throw (PAYT) policies. Such programs aim to internalize externalities of municipal waste pollution by pricing unsorted waste per unit. This paper uses a multiple random forest approach on a unique six-year panel data of waste generation and a large set of covariates for about all Italian municipalities. Probability of PAYT adoption is found to be heterogeneous, and positively correlated with pre-policy recycling rates, distance to the nearest incinerator and to municipalities that already implemented the policy. PAYT results to be effective, decreasing per capita unsorted waste by 30% on average. This effect seems mostly driven by waste avoidance (−8%), and only to a lesser extent by an increase in recycling. Effects for unsorted and total waste are dynamic, leading to lower waste levels in the long-run. PAYT effectiveness on unsorted and total waste is found heterogeneous in pre-policy waste levels, housing density, voters' political inclination, and education level. Policy-compliers are estimated to have low housing density, low pre-policy unsorted waste and high recycling, suggesting a positive effect of environmental awareness on PAYT effectiveness.

Keywords: Waste generation, Pay-as-you-throw, Heterogeneous treatment effects, Random forests
JEL Codes: C14, C21, C52, Q53

*Humboldt University of Berlin and DIW Berlin – German Institute for Economic Research, Berlin, Germany. Email: mvalente@diw.de

1 Introduction

In the presence of pollution externalities, environmental policy aims to internalize the difference between social and private costs into individuals' polluting behavior. Empirical evaluations of such policies can help policy makers and local governments to understand which factors may facilitate policy implementation, determine policy effectiveness, and safeguard against adverse policy outcomes. Municipal Solid Waste (MSW) generation characterizes a classic example of negative externality, as its environmental and treatment costs are higher than its private ones. Standard flat fees on MSW collection, however, do not usually suffice for internalizing the cost difference into individuals' waste generation behavior. Therefore, charging households on the basis of the actual amount of waste generated, a form of Pigouvian fee (Pigou, 1932), has been increasingly used as a policy instrument known as unit pricing or Pay-As-You-Throw (PAYT) program. PAYT provides direct incentives to reduce unsorted waste, the priced waste amount, and indirect ones to increase recycling, usually left unpriced, and to decrease total waste generation, with the latter being the most desirable outcome from a social viewpoint.[1]

Using a unique and large dataset spanning from 2010 to 2015 for about all Italian municipalities, this paper looks at the way households change waste generation behavior in response to price incentives by estimating heterogeneous and dynamic causal effects of PAYT policies in Italy on per capita unsorted, recycling, and total waste. To the best of my knowledge, this is the first paper that analyzes policy adoption and effectiveness including a large set of municipal characteristics, and estimates PAYT effects using non-parametric machine learning methods. Thereby, this research also aims at advising non-PAYT municipalities by predicting waste pollution savings for the policy compliers, i.e., those non-PAYT municipalities that, based on their characteristics, are most likely to

[1] Early discussions on PAYT effectiveness can be found, e.g., in the New York Times (Hanley, 1992) and later in, e.g., the European Union's Directives on waste (European Union, 2018).

reduce waste and increase recycling if they were to introduce PAYT.

Even though the literature generally estimates negative average effects on unsorted waste, there is no consensual evidence on the behavioral mechanisms behind such reduction, i.e., on whether households adjust their behavior by increasing recycling and decreasing total waste. Existing results vary depending on the method, data and municipalities under study.[2] This may be due to several reasons. In particular, available data is often small in either time or cross-sectional dimension or both, which questions its representativeness and results' external validity (Allers & Hoeben, 2010). Further, the set of control variables is usually limited, not allowing to include all relevant municipal characteristics explaining waste generation and, if non-random, policy adoption. As a result, omitted variable bias and policy endogeneity threaten the internal validity of existing studies. Additionally, parametric model assumptions limit model flexibility and may lead to misspecification. In sum, to provide further evidence on the effectiveness of waste policies, it seems necessary to analyze possible (observed) determinants of waste generation, policy adoption and heterogeneous effects, estimating them in a more flexible framework.

My empirical strategy uses the potential outcome approach (Rubin, 1974) to identify heterogeneous policy (treatment) effects under unconfoundedness and no spillover (SUTVA) assumptions[3]. Estimation proceeds via residualization to account for non-random policy adoption, and the potential impact of covariates on both waste generation and adoption (Neyman, 1979; Robinson, 1988). Residualization amounts to, first, partial out the effect of the covariates on outcomes and treatment propensities and, second,

[2]For example, waste reductions of variable magnitude but no statistically significant effects on recycling are found, for the U.S., in Jenkins *et al.* (2003) and Fullerton & Kinnaman (2000, 1996): for the Netherlands, in Allers & Hoeben (2010), and in Valente & Bueno (2018) for Italy. Instead, Carattini *et al.* (2018) and Bucciol *et al.* (2015) find a positive effect on the frequency of recycling for Switzerland, and on recycling rates for Italy, respectively. For Korea, Hong & Adams (1999) estimate positive effects recycling, but no source-reductions. For Japan, Usui & Takeuchi (2014) find long-run increases in recycling and only short-run total waste reductions.

[3]I.e., policy adoption is as good as random once controlled for the covariates, and potential outcomes of one unit are unaffected by policy adoption by other units, respectively.

predict the residualized outcomes with the residualized propensity scores. Heterogeneous conditional means and treatment effects are estimated by random forests (RFs) (Athey *et al.* , 2018)[4]. RFs are non-parametric and data-driven conditional mean estimators employed in this paper with the aim to avoid possible model misspecification; include a large set of (often correlated) covariates and decorrelate their partial effects by means of randomization; estimate heterogeneous policy effects for each municipality, and perform policy targeting, i.e., predict policy effects for non-PAYT municipalities based on their characteristics.

Results are still preliminary, and are summarized in the last Section (6) of this paper.

The remainder of the paper is organized as follows: Section 2 provides a brief review of the related literature on PAYT, and discusses potential methodological shortcomings. Section 3 outlines the random forest approach used to estimate the policy effects under study. Background and data are described in Section 4. Section 5 shows the results, and Section 6 concludes and provides policy implications of this study.

2 Literature

3 Method

To empirically learn not only the magnitude of policy effects but also the determinants of policy adoption and effectiveness, I estimate policy heterogeneous effects as a function of a large set of municipal characteristics that may influence policy adoption and responsiveness. Following the potential outcome approach (Rubin, 1974), let (X_i, Y_i, D_i) be the available data for municipality $i = 1, ..., n$, where $X_i = x \in \mathbb{R}^p$ is a vector of p covariates, Y_i is the waste outcome, and $D_i \in \{1, 0\}$ is the policy (treatment) dummy under {PAYT,

[4]This method builds upon, e.g., Athey & Wager (2018); Athey & Imbens (2016); Breiman (2001).

non-PAYT}. Additionally, let $\{Y_i(1), Y_i(0)\}$ be the waste outcomes that we would observe under $D_i = 1$ or 0 respectively, such that $Y_i = Y_i(D_i)$. For each municipality i defined by its vector of characteristics x, the goal is to estimate the conditional Treatment Effect (TE) $\delta(x) = \mathbb{E}\left[Y_i(1) - Y_i(0)|X_i = x\right]$ under unconfoundedness,[5], and the no spillover assumption, aka SUTVA (Rubin, 1974). To fulfill unconfoundedness in the likely case of non-random PAYT assignment, it is necessary to control for the potential sources of self-selection into policy, i.e., capture the effect of X_i on D_i. Consider the waste outcome model:

$$Y_i = \mathbb{E}\left[Y_i(0)|X_i\right] + D_i\delta(X_i) + \epsilon_i, \tag{1}$$

where $\mathbb{E}\left[\epsilon_i|X_i, D_i\right] = 0$ holds under unconfoudedness. Let $\mathbb{E}\left[Y_i|X_i\right]$ and $\mathbb{E}\left[D_i|X_i\right]$ be the conditional means of the waste outcome Y_i and policy adoption D_i respectively, with $\mathbb{E}\left[D_i|X_i\right]$ representing policy adoption probabilities aka propensity scores. After few algebraic transformations, model (1) can be rewritten as:

$$Y_i - \mathbb{E}\left[Y_i|X_i\right] = (D_i - \mathbb{E}\left[D_i|X_i\right])\delta(X_i) + \epsilon_i. \tag{2}$$

Thereby, estimation of the treatment effect $\delta(X_i)$ in (2) proceeds via residualization of the conditional means $\mathbb{E}\left[Y_i|X_i\right]$ and $\mathbb{E}\left[D_i|X_i\right]$ which are separately estimated in a previous step (Neyman, 1979; Robinson, 1988). In particular, estimating $\mathbb{E}\left[D_i|X_i\right]$, namely, the selection equation, and $\delta(X_i)$, namely, heterogeneous policy effects by observed covariates, allows to analyze the determinants of policy adoption (self-selection) and effectiveness, respectively. Since there is no prior information on the functional form of these possibly high-dimensional relations, employing a parsimonious parametric model could not possibly capture complex interactions or nonlinear terms. Also, including a large set of potential predictors in such a model may lead to, e.g., variance inflation and incorrect signs. Thus,

[5] Namely, $\{Y_i(0), Y_i(1)\} \perp\!\!\!\perp D_i|X_i$. This assumption is plausible as we include a large set of covariates with a potential impact on outcome levels, policy adoption, and policy effectiveness.

to reduce the risk of model misspecification and ad-hoc model selection, these functions are determined non-parametrically using RFs (Athey & Wager, 2018; Athey & Imbens, 2016; Breiman, 2001) and, specifically, the generalized version of this estimator developed in Athey *et al.* (2018).

RFs are ensamble of trees, and each tree is a non-parametric conditional mean estimator used to estimate the quantities of interest defined as $y(x) = \mathbb{E}[Y_i|X_i = x]$, $d(x) = \mathbb{E}[D_i|X_i = x]$ and $\delta(x)$. Such estimator works according to the principle "divide and rule": Every tree recursively partitions the covariate space into binary regions according to an optimal criterion. In this case, such criterion is chosen to maximize outcome heterogeneity between regions. Thereby, estimation of a tree, namely, a sequence of binary regions, proceeds greedily, i.e., partitions that reduce outcome heterogeneity the most are performed first. As a result, final regions aka leaves contain a number of observations, set to be greater than a given minimum,[6] that are as homogeneous as possible in terms of those covariates able to explain most of the variation in the quantity of interest. As a result, the estimated conditional expectation function is a nearest-neighbor function that matches units with similar characteristics by partitioning them into optimal regions, and estimates the conditional mean of interest within each final region. In addition, each tree in a RF is built using a random subsample of data and covariates, aka training sample. The estimated variance from one single tree is usually large because each tree is built to minimize bias. In this respect, RFs avoid overfitting, i.e., reduce such variance, by averaging estimates over trees. Once the RF is estimated (trained), unbiased conditional mean estimates for each vector $X_i = x$ are obtained via out-of-bag prediction. This amounts to obtain (predict) the quantity of interest for $X_i = x$ from trees (bags) estimated without it.[7]

[6]Note that minimum leaf-size and other parameters such as a penalty for imbalanced partitions are chosen via cross-validation according to Tibshirani *et al.* (2018).

[7]Without out-of-bag prediction, random forest estimators are asymptotically normal, since each estimate is derived by averaging estimates from many trees, but they overfit and do not converge at the

Formally, for each i in a random subsample of training data s^{tr}, define $\theta(x)$ the quantity to estimate, namely, $y(x), d(x)$, and $\delta(x)$. As defined above, these quantities are conditional expectations using outcome(s) $Q_i = \{Y_i, D_i\}$. A tree recursively partitions the covariate space into binary regions such as to maximize $\hat{\theta}$-heterogeneity between regions (Athey *et al.* , 2018):

$$\frac{N_{R_1} N_{R_2}}{N_{R_0}^2} (\hat{\theta}_{R_1}(s^{tr}) - \hat{\theta}_{R_2}(s^{tr}))^2, \tag{3}$$

with $\frac{N_{R_{1,2}}}{N_{R_0}}$ being the fraction of training examples $i : X_i \in s^{tr}$ assigned to two new regions $R_{1,2}$ from the original region R_0. Such new regions are defined by optimal covariate cutoffs that maximize (3) greedily. As in Athey *et al.* (2018), $\hat{\theta}_{R_j}(s^{tr})$ is identified by the local estimating equation $\mathbb{E}\left[\rho_\theta(Q_i)|X_i = x\right] = 0$ where ρ_θ is the moment function. The solution to this estimating equation is defined as $\hat{\theta}_{R_j}(s^{tr})$, and it is estimated as follows:[8]

$$\hat{\theta}_{Rj}(s^{tr}) \in \underset{\theta}{\operatorname{argmin}} \|\sum\nolimits_{i:X_i \in R_j} \rho_\theta(Q_i)\|_2. \tag{4}$$

The moment function for the conditional mean estimation of $y(x)$ and $d(x)$ writes $\rho_\theta(Q_i) = Q_i - \theta(x)$, with $Q_i = \{Y_i, D_i\}$ respectively. In this case, equation (4) gives the standard random forest's solution $\hat{\theta}(x) = \bar{Q}$ from Breiman (2001), which is the mean outcome computed over $\forall i : X_i \in R_j$ of the b-th tree and then averaged over all B trees ($b = 1, ..., B$). Predictions for each $X_i = x$ are obtained out-of-bag, and $\{\hat{y}(x), \hat{d}(x)\}$ are residualized as $\{\tilde{Y}_i = Y_i - \hat{y}(x), \tilde{D}_i = D_i - \hat{d}(x)\}$. Similarly, the moment function for treatment effect estimation of $\delta(x)$ uses the residualized outcomes, and writes $\rho_{\theta,\alpha}(\tilde{Y}_i, \tilde{D}_i) = (\tilde{Y}_i - \delta(x)\tilde{D}_i - \alpha(x))(1\ \tilde{D}_i)^\top$ where $\alpha(x)$ is an intercept term. Further, the TE is identified by the local moment con-

square-root-n rate, thus, are bias-dominated (Athey & Imbens, 2016; Mentch & Hooker, 2016).

[8]Due to computational inefficiency, Athey *et al.* (2018) optimize a gradient-based approximation to (3) and (4) requiring to evaluate $\hat{\theta}_{R_0}$ and the moment function only in R_0, and not in each new region $R_{1,2}$ (see Athey *et al.* , 2018, for more details).

dition $\mathbb{E}\left[\rho_{\theta,\alpha}(\tilde{Y}_i,\tilde{D}_i)|X_i=x\right]=0$ as $\delta(x)=Var[\tilde{D}_i|X_i=x]^{-1}Cov[\tilde{D}_i,\tilde{Y}_i|X_i=x]$. Using $\rho_{\theta,\alpha}(\tilde{Y}_i,\tilde{D}_i)$, each tree is estimated as in (4).[8] For out-of-bag prediction, forest-based weights $w_i(x)$ are computed $\forall X_i=x$ as the frequency with which the training example $i: X_i \in s^{tr}$ is estimated in the same region as the held-out x. This is meant to correct for randomization bias, e.g., noisy leaves.[9] Therefore, the solution to the (weighted) estimating equation for the TE yields:

$$\hat{\delta}(x)=(\sum_{i=1}^{n}w_i(x)(\tilde{D}_i-\bar{D}_w)^2)^{-1}\sum_{i=1}^{n}w_i(x)(\tilde{D}_i-\bar{D}_w)(\tilde{Y}_i-\bar{Y}_w) \tag{5}$$

where $\bar{D}_w=\sum_{i=1}^{n}w_i(x)\tilde{D}_i$ and $\bar{Y}_w=\sum_{i=1}^{n}w_i(x)\tilde{Y}_i$.

Finally, the variance of $\hat{\theta}(x)$ is estimated by evaluating the estimator on bootstrapped half-samples of the training data, aka bootstrap of little bags (see, for a discussion, Athey *et al.* 2018, and Sexton & Laake 2009).

4 Background and Data

4.1 PAYT policies

PAYT policies in Italy, as in many municipalities worldwide, require households to pay a fix price per unit of unsorted waste according to either its volume, i.e., per bag or can of waste, or weight, i.e., per kilogram of waste. The baseline policy in both PAYT and non-PAYT municipalities is a flat fee independent of waste quantities, namely, the unit price is zero. Flat fees depend on household size and the floor area of the house. Municipalities are cost minimizers and set the flat fee to cover fixed and variable costs of waste collection. PAYT municipalities implement PAYT on top of the flat fee, and reduce the latter because PAYT covers variable costs of waste collection. Both flat and PAYT fees

[9] See Athey *et al.* (2018) for the formal definition of such forest-based weights.

are adjusted yearly based on waste amounts and costs of the previous year. Enforcement and monitoring by both the waste collection company and the municipal police are in place to prevent policy adverse effects.[10]

At the time of writing, there is no database including information on the type of PAYT as well as the unit price for each municipality. Thereby, I cannot evaluate PAYT effects separately for each system, as well as compute price elasticities. However, considering a sample of PAYT municipalities over the years reveals that the variation of PAYT prices is limited, varying between 0.05 to 0.15 euros per liter of unsorted waste. Moreover, an assumption that I must make comes from the fact that in many PAYT municipalities the unit price for the first q bags or cans of waste is zero. Such q varies between municipalities, and yearly adjusted as for the unit price. In this respect, I assume that every household generates more than q bags or cans, thus, faces a non-zero unit price.[11]

4.2 Outcomes

I collected data for about all Italian municipalities over the years 2010-2015. Waste data come from the database of the Italian Institute for Environmental Protection and Research (ISPRA) which also provided information about PAYT introduction for some municipalities. Full data on PAYT adoption were collected upon direct request, and provided by waste collection companies and the National Association of Italian Municipalities (ANCI). The outcome of the selection model is a binary variable that takes the value of one if the municipality introduced PAYT and zero otherwise. Outcomes of the treatment effect model are kilograms of per capita unsorted (UW), recycling waste (RW), and total (TW) which is the sum of the previous two. I also estimate effects on the recycling rate

[10] E.g., trash bins are locked, waste collectors do not remove bags or cans if sensed to weigh excessively (under volume-based PAYT), all households in a building and/or single households are fined if unsorted waste is found inside the recycling and in case of illegal dumping.

[11] Yet, I am in the process of creating a database on PAYT types, unit prices, and q values for all PAYT municipalities.

($RWrate$) for comparison. Figure 1 provides information of the geographical distribution of per capita waste amounts at municipality level. In general, northern municipalities generate higher amounts of total waste per capita but also higher recycling and recycling rate.

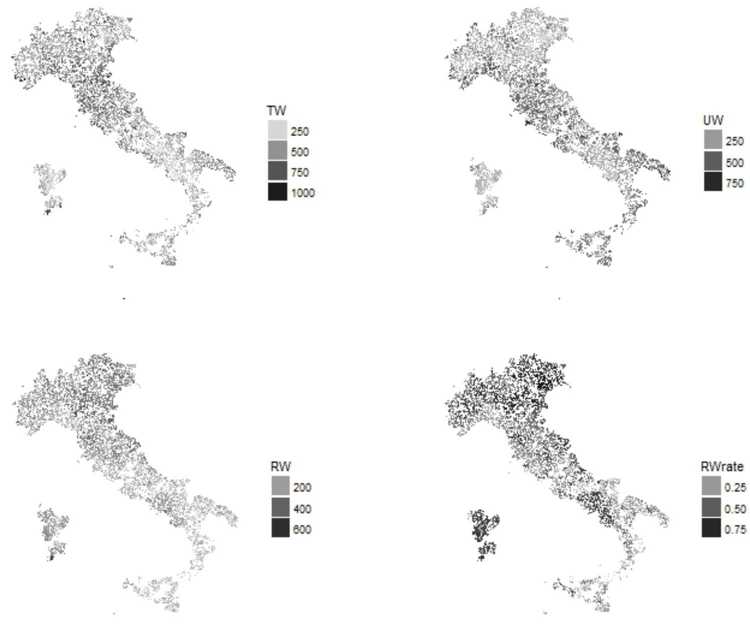

Figure 1: Average per capita waste amounts by municipality over 2010-2015 (white, no data) - unsorted (UW), total (TW), recycling (RW) waste and rate ($RWrate$).

I consider as treated units those municipalities that implemented PAYT from 2012 onwards, and discard PAYT municipalities before 2012. This choice is due to the fact that lagged waste amounts are used as predictors in all equations (selection, waste generation, and treatment effect), and that first-order lags for PAYT municipalities are likely biased predictors. This may be due to two reasons. First, as common in PAYT policy settings, households likely anticipate policy in the year before treatment and vary waste generation behaviors. Second, PAYT municipalities usually introduce a training period to the new system before its official start in which waste is generated under unit pricing but it is not priced. As this cannot be observed by the researcher and its potential effects are unknown, I use second- and not first-order lags of waste generation for the treated units to predict policy adoption, and policy effects in the first treatment year. Also, such first-order lags can be used to estimate policy (anticipation) effects. The resulting treated municipalities are 155, of which 96 in the North-West, 54 in the North-East, and 5 in the Center. Additionally, since mismeasurement of waste data for the South is more likely due to illegal waste disposal, I exclude all observations in the South from the data.[12] Further, most of the treated municipalities implement PAYT for the first time in 2013 (69), while the others in 2012 (20), 2014 (41), and 2015 (25). I analyze dynamic treatment effects by running separate RFs for the first, second, third, and fourth year from treatment as well as the year before official policy start (anticipation). Figure 2 shows the distribution of PAYT (black) and non-PAYT (grey) municipalities over time.

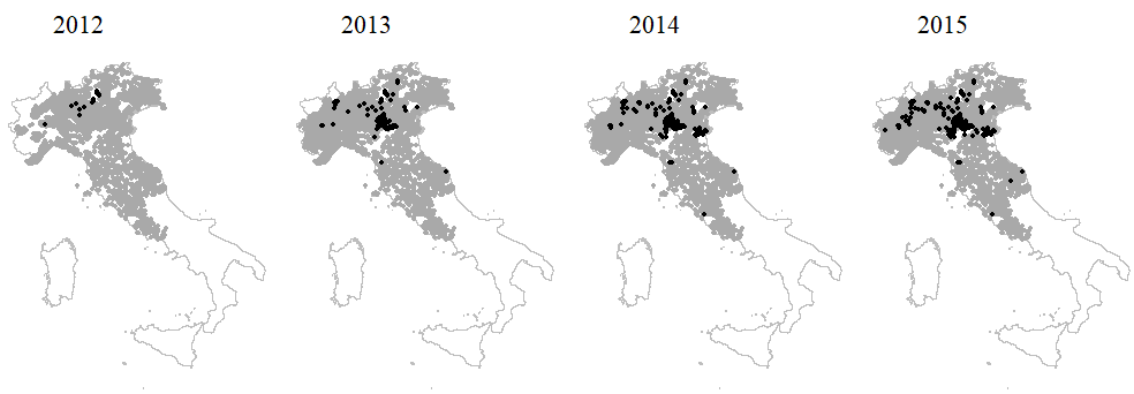

Figure 2: Map of PAYT (black) and non-PAYT (grey) municipalities over 2010-2015.

For a descriptive comparison of waste generation at the beginning (2010) and at the end (2015) of the sample, Figure 3 plots the distribution of UW, TW, and RW for PAYT (grey) and non-PAYT (white) municipalities.

[12]Between 2010 and 2015, average crime rates on illegal incineration and traffic of waste and violations of waste management regulations including, e.g., illegal dumping, amount to 29% and 14%, respectively, for the South compared to a national average of 10% for both crimes (ISTAT, 2018).

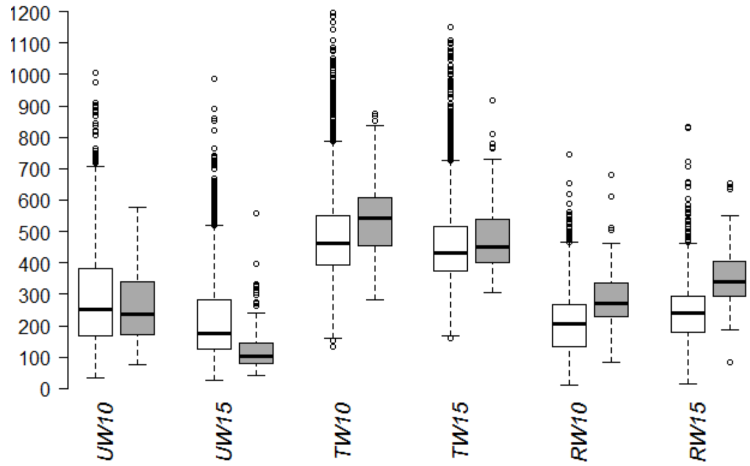

Figure 3: Boxplot of 2010 and 2015 per capita unsorted (UW), total (TW), recycling (RW) waste for PAYT (grey) and non-PAYT (white) municipalities.

In both PAYT and non-PAYT municipalities we observe a decrease in UW and TW, and an increase in RW over time. In particular, the decrease in UW for PAYT municipalities is substantial. Further, comparing pre-policy (2010) waste amounts in PAYT and non-PAYT municipalities reveals possible self-selection into policy. Figure 3 shows that PAYT municipalities present relatively higher TW and RW levels. The first may suggest that municipalities introduced PAYT due to high waste pollution and collection/treatment costs. The second indicates that municipalities introducing PAYT have already developed recycling habits before adoption. Thereby, to minimize the risk of adverse effects such as illegal dumping, PAYT may be implemented only if household opportunity costs of recycling are already low before policy.

4.3 Determinants of PAYT adoption and waste generation

As covariates, I use waste generation and policy adoption determinants for a total of 46 variables. We include socio-economic variables typically chosen in the literature[13] which

[13] See, e.g., Grossmann *et al.* (1974); Jenkins *et al.* (2003); Miranda & Bauer (1996); Richardson & Havlicek (1978); Van Houtven & Morris (1999), and Wertz (1976).

are, e.g., average household size (*hhSize*); per capita income (*incomePc*); educational attainment as measured by the share of the population with a graduate degree or higher (*collegeDeg*) as well as with an elementary degree or none (*elemDeg*); tourism intensity, measured as the nights spent by tourists divided by the local population (*tourism*); and age structure, decomposed into citizens under 15 (*age14*) and over 65 years old (*age65*). Moreover, I control for labor market characteristics such as the share of unemployed and out-of-the-labor force population (*unempOutLab*), and an index of the labor market activity in terms of commuting intensity (*labMarket*) which may determine the time spent at home and the waste generated. Further, I account for observables that possibly serve as fixed effects, such as the degree of urbanization (*urban*, *highUrban*), and the type of municipality, i.e., regional and provincial seats (*regSeat*, *provSeat*). Moreover, I also include the distance of each municipality to incinerator (*distInc*), landfill (*distLandf*), and hazardous waste treatment facilities (*distHaz*) because this may impact not only waste generation incentives, but also policy adoption. In fact, the latter may be more likely in a municipality where the municipal landfill is full and the incinerator is far away. However, the opposite may also occur because the presence of a polluting incinerator in the vicinity may stimulate households' environmental awareness and lead to PAYT adoption to limit environmental and health damages. Importantly, to control for possible policy interaction as PAYT municipalities seem clustering in certain areas (see Figure **??** in the Appendix), I also include the distance to each non-PAYT municipality in a given year to the closest municipality that implemented PAYT in earlier years. This should control for information dissemination, thus, a potentially higher probability of policy adoption for municipalities that mimic successful neighbors' policies (Allers & Hoeben, 2010). The last group of covariates includes a proxy for political participation (*polPart*), and political preferences (*votesLeft*, *votesRight*, *votesBigTent*), as well as mayors' characteristics such as age (*mayorAge*), term length (*yearsOffice*), and political party (e.g., *mayorGreen*).

Finally, as mentioned above, lagged waste amounts are used as predictors that likely explain policy adoption and effectiveness, as well as other unobserved, persistent waste generation habits (*UWpre*, *RWpre*, *TWpre*, *RWratePre*). In particular, lagged recycling controls for existing differences in recycling systems, e.g., curbside recycling, that may favor sorting habits. Variables' descriptions and descriptive statistics for PAYT and non-PAYT municipalities are presented in the Appendix 6.

Socio-economic, tourism, and geographic data is collected from the Italian National Institute of Statistics (ISTAT) and through webscraping the online database of comuni-italiani.it. Data on municipal political elections and mayors' characteristics are obtained from the Ministry of the Interior upon request. The geographic location of landfills, incinerators and other waste treatment facilities is computed with data of the European Pollution Release and Transfer Register (E-PRTR).

5 Results

5.1 Event study

For a conditional comparison of waste generation between treated and control group that accounts for fixed differences between municipalities, Figure 4 presents the results from an event-study estimation. This corresponds to a Difference-in-Differences (DiD) regression including lead and lagged policy dummies for PAYT municipalities equal to one in each pre- and post- policy year, respectively. As a result, I include four of such leads and lags using the last lead before policy as a baseline dummy. The aim of such DiD is to, first, give an indication of the magnitude of average policy effects in the short- and in the long-run, and, second, to test for differences in waste generation trends between PAYT and non-PAYT units pre-policy, which would cause biased DiD estimates and suggest self-selection into policy.

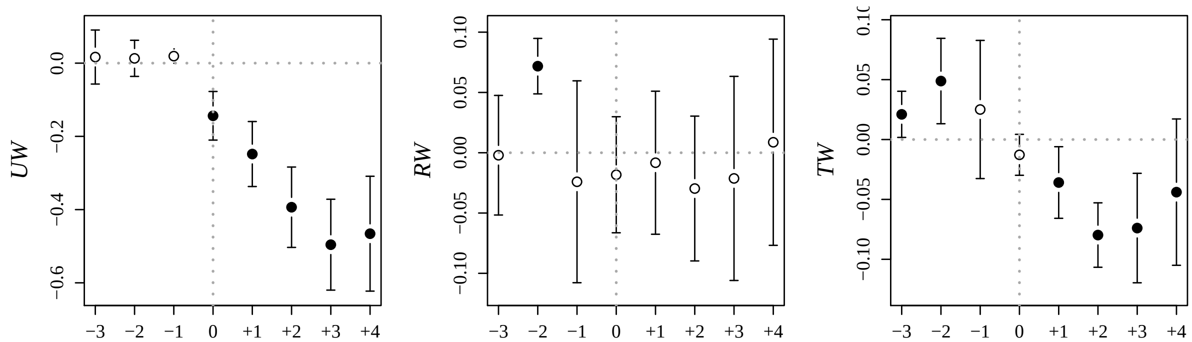

Figure 4: DiD estimates (2010-2015) of lead and lagged average effects on log p.c. waste with municipal fixed effects and Driscoll & Kraay (1994) robust standard errors. Black dots (vertical dotted lines) represent statistical significant effects at 5% (policy decison year).

Figure 4 plots estimated lead and lagged average effects with their confidence intervals. Black dots represent statistical significant effects at 5%, and vertical dotted lines indicate the year of policy decision which usually includes PAYT training, namely, households discard their waste per bag/weight/emptying but do not pay accordingly. Post-policy (lagged) estimates show that PAYT is effective in reducing UW and TW especially until the third year (+3) after policy decision with estimates in the range of 15-50% (4-8%) for UW (TW). However, such policy effects reduce in magnitudes in the fourth (+4) year after policy. No statistical significant policy effects are found on recycling. In light of this, exploring PAYT heterogeneous effects seems crucial to understand determinants of policy compliance. Further, considering the pre-policy period, leads are statistically significant in some cases, indicating that PAYT-municipalities have on average higher RW and TW pre-policy (as suggested also in Figure 3). This reveals that waste generation trends between PAYT and non-PAYT municipalities differ before policy, likely because policy adoption determinants are not controlled for. In this case, DiD estimates of post-PAYT effects are biased due to violation of the identifying parallel trends assumption (PTA). For example, waste generation in PAYT municipalities may follow unit-specific trends already before policy because, e.g., PAYT may be adopted by municipalities with ceteris paribus higher per capita recycling or total waste (or both). Also, DiD only allows for unit effects that have time-invariant effects on the outcome, which might not be the case in this setting. If, e.g., environmental-friendly attitudes or time and space availability for recycling have evolving effects over time, then fixed effects cannot capture these time-varying unit effects. Controlling for this variation with a large set of covariates explaining both policy adoption and waste generation is at the heart of this paper's motivation to employ a RF approach.

5.2 Random forests

[Preliminary results are summarized in the Conclusions below.]

6 Conclusions

This paper shows that PAYT is effective, reducing per capita (pc) unsorted waste amounts by 30% on average in the first policy year. This effect seems to be largely driven by behavioral changes towards waste avoidance (−8%) and possibly by a smaller increase in recycling (+6%, though not statistically significant). This means that households mostly respond to the policy adopting source-reduction behaviors such as reuse. By comparing these results to those obtained by a difference-in-differences approach, I show that failing to account for time-varying, municipal-specific determinants of waste generation likely driving self-selection into policy may lead to a mismeasurement of policy effects due to PTA violation.

Using a random forest approach, probability of PAYT adoption is found to be heterogeneous depending on, e.g., pre-policy recycling rates, the distance to the nearest incinerator, other PAYT municipalities that already implemented the policy, tourism intensity, education level as well as political variables such as age and political party of the mayor.

In addition, PAYT average effects for unsorted and total waste are dynamic, increasing by on average 3% in every year after the first implementation year. This indicates that households adjust and improve waste generation behavior over time. On the contrary, average dynamic effects on recycling are rather stable over time and are never found to be statistically significant.

Analyzing PAYT effects by observed municipal characteristics, I find heterogeneity in policy effectiveness depending on, especially, pre-policy waste levels, housing density, voters' political inclination, and education level. Although PAYT policies do not seem successful to stimulate recycling on average, estimated heterogeneous policy effects reveal that PAYT significantly increases recycling only in those municipalities with relatively low pre-policy recycling pc. This points towards positive effects on recycling only in the presence of large marginal improvements in sorting behaviors before policy. Further, I estimate PAYT effectiveness on unsorted and total waste to be higher for highly educated municipalities with low housing density as well as relatively low pre-policy unsorted waste and high recycling pc, suggesting a positive effect of environmental-awareness on PAYT effectiveness.

Results on waste pollution savings for non-PAYT municipalities and policy compliers are still preliminary, and will be presented at your Conference in case of admission.

This study has several potential policy implications. Adopting PAYT seems to lead to waste pollution reduction and changes in household waste generation behavior. However, PAYT effects are dynamic and heterogeneous, i.e., depend on time from adoption and on municipal characteristics, respectively. Accounting for this, municipalities could increase policy effectiveness by, e.g., providing incentives for sorting and source-reduction behaviors already before policy implementation.

Appendix

Table 1: Descriptives for PAYT (P) and non-PAYT (N) municipalities (2010-2015)

	mean(P)	min(P)	max(P)	SD(P)	mean(N)	min(N)	max(N)	SD(N)
RW	324.88	82.79	655.20	92.28	225.08	10.02	835.03	91.37
TW	449.70	146.98	920.07	116.35	460.19	114.54	1193.25	129.85
UW	124.82	29.99	584.12	78.33	235.11	23.37	985.09	133.77
RWrate	0.73	0.18	0.89	0.12	0.51	0.02	0.91	0.19
RWratePre	0.68	0.18	0.89	0.14	0.48	0.02	0.90	0.20
RWpre	307.05	82.79	612.02	86.63	217.02	10.02	819.72	92.52
TWpre	461.68	146.98	920.07	122.12	466.25	134.24	1195.52	131.92
UWpre	154.63	29.99	590.88	97.52	249.23	23.37	1004.85	138.12
distLandf	7.64	0.00	23.67	6.16	12.11	0.00	89.93	11.54
distHaz	9.28	0.00	48.13	8.63	15.10	0.00	115.40	16.31
distInc	24.82	0.00	72.49	18.74	31.82	0.00	178.01	32.25
distPayt	11.74	0.00	326.11	25.23	67.29	1.00	371.26	87.13
migrNet	0.00	-0.07	0.07	0.01	0.00	-0.13	0.15	0.01
pop	8.78	0.12	192.84	18.17	7.12	0.03	1345.85	33.27
hhSize	2.28	1.71	3.00	0.29	2.26	1.00	4.92	0.29
popDens	2.73	0.06	23.81	4.01	3.18	0.01	77.66	5.58
popGrowth	-0.01	-0.72	0.56	0.13	-0.01	-2.44	1.35	0.16
income	105.52	7.29	22197.00	1325.40	13.72	2.96	34.61	2.40
males	0.49	0.41	0.54	0.01	0.49	0.41	0.69	0.02
foreignPop	0.10	0.01	0.19	0.04	0.08	0.00	0.41	0.04
tourism	0.34	0.00	8.94	0.97	0.51	0.00	53.26	1.60
age14	0.14	0.09	0.21	0.02	0.13	0.01	0.22	0.03
age65	0.22	0.12	0.33	0.04	0.23	0.05	0.64	0.06
age0	0.05	0.03	0.06	0.01	0.04	0.03	0.06	0.01
elemDeg	0.31	0.20	0.42	0.04	0.31	0.11	0.58	0.05
collegeDeg	0.09	0.03	0.17	0.02	0.09	0.03	0.18	0.03
oneParentFam	9.69	6.26	14.58	1.31	9.92	0.00	19.44	1.80
rentedHouses	8.84	4.13	15.12	1.93	9.40	0.00	44.12	3.15
hhPerSqMeter	2.21	1.68	3.08	0.26	2.31	1.19	3.34	0.28
deprIndex	-1.94	-6.42	1.11	1.26	-1.26	-7.62	7.98	1.72
outLabRate	0.62	0.51	0.78	0.05	0.64	0.46	0.90	0.06
unempOutLab	6.12	0.87	11.98	1.65	6.86	0.00	24.00	2.72
labMarket	0.39	0.11	0.66	0.11	0.42	0.00	0.66	0.14
polPart	0.70	0.49	0.82	0.07	0.69	0.30	0.91	0.08
votesLeft	0.06	0.02	0.24	0.03	0.07	0.00	0.84	0.09
votesRight	0.12	0.03	0.31	0.06	0.13	0.00	0.57	0.08
votesBigTent	0.23	0.12	0.37	0.05	0.24	0.00	0.58	0.07
mayorAge	51.14	21.00	78.00	10.88	52.17	22.00	84.00	10.59
yearsOffice	1.82	0.00	5.00	1.57	1.86	0.00	5.00	1.40
mayorLeft	0.09	0.00	1.00	0.29	0.06	0.00	1.00	0.23
mayorRight	0.05	0.00	1.00	0.22	0.09	0.00	1.00	0.28
mayorGreen	0.03	0.00	1.00	0.18	0.02	0.00	1.00	0.12
mayorOther	0.18	0.00	1.00	0.38	0.10	0.00	1.00	0.29
mayorReg	0.64	0.00	1.00	0.48	0.74	0.00	1.00	0.44
localMayor	0.79	0.00	1.00	0.41	0.79	0.00	1.00	0.41
highUrban	0.10	0.00	1.00	0.30	0.14	0.00	1.00	0.34
urban	0.59	0.00	1.00	0.49	0.42	0.00	1.00	0.49
provSeat	0.02	0.00	1.00	0.15	0.01	0.00	1.00	0.10
regionSeat	0.01	0.00	1.00	0.09	0.00	0.00	1.00	0.05
noSeat	0.97	0.00	1.00	0.18	0.99	0.00	1.00	0.11

Table 2: Variables' description. Census indicates 2011 values (ISTAT, 2011).

	Variables' description
RW	recycling waste per capita (kg)
TW	total waste per capita (kg)
UW	unsorted waste per capita (kg)
RWrate	recycling rate (% of total waste)
RWratePre	lagged recycling rate (% of total waste)
RWpre	lagged recycling waste per capita (kg)
TWpre	lagged total waste per capita (kg)
UWpre	lagged unsorted waste per capita (kg)
distLandf	distance to closest waste landfill (km)
distHaz	distance to closest hazardous waste treatment facility (km)
distInc	distance to closest waste incinerator (km)
distPayt	distance to closest PAYT municipality in t-1 (km)
migrNet	net migration flow per capita
pop	population (x 1,000 inhabitants)
hhSize	average household size (n. household members)
popDens	population density (inhabitants per km2)
popGrowth	population growth
income	income per capita ((x 1,000 euros)
males	share of male population
foreignPop	share of foreign population
tourism	nights spent by tourists per capita (x 1,000)
age14	share of population aged less than 14
age65	share of population aged more than 65
age0	share of population aged less than 5 (census)
elemDeg	share of population with elementary degree or lower (census)
collegeDeg	share of population with college degree (census)
oneParentFam	share of single-parent families (census)
rentedHouses	share of rented houses (census)
hhPerSqMeter	housing density (inhabitants per m2, census)
deprIndex	social deprivation index (census)
outLabRate	out-of-the-labor force rate (census)
unempOutLab	unemployed and out-of-the-labor force rate (census)
labMarket	commuting intensity index (IIRFL index, census)
polPart	voter turnout in the 2013 Italian general election (IGR13)
votesLeft	vote shares of left-wing parties in the IGR13
votesRight	vote shares of right-wing parties in the IGR13
votesBigTent	vote shares of big-tent parties in the IGR13
mayorAge	mayor's age
yearsOffice	mayor's term of office (years)
mayorLeft	left-wing mayor (dummy)
mayorRight	right-wing mayor (dummy)
mayorGreen	green-party mayor (dummy)
mayorOther	mayor of other party (dummy)
mayorReg	mayor of local party (dummy)
localMayor	mayor born in the municipal province (dummy)
highUrban	highly urbanized municipality (dummy)
urban	medium urbanized municipality (dummy)
provSeat	provincial capital (dummy)
regionSeat	regional capital (dummy)
noSeat	no capital (dummy)

References

Allers, M., & Hoeben, C. 2010. Effects of unit-based garbage pricing: A differences-in-differences approach. *Environmental Resource Economics*, **45**(3), 405–28.

Athey, S., & Imbens, G. 2016. Recursive Partitioning for Heterogeneous Causal Effects. *PNAS*, **113**(27), 7353–7360.

Athey, S., & Wager, S. 2018. Estimation and Inference of Heterogeneous Treatment Effects using Random Forests. *Journal of the American Statistical Association*, **113**(523), 1228–1242.

Athey, S., Tibshirani, J., & Wager, S. 2018. Generalized Random Forests. *Forthcoming, Annals of Statistics*.

Breiman, Leo. 2001. Random Forests. *Machine Learning*, **45**(1), 5–32.

Bucciol, A., Montinari, N., & Piovesan, M. 2015. Do Not Trash the Incentive! Monetary Incentives and Waste Sorting. *The Scandinavian Journal of Economics*, **117**(4), 1204–1229.

Carattini, S., Baranzini, A., & Lalive, R. 2018. Is Taxing Waste a Waste of Time? Evidence from a Supreme Court Decision. *Ecological Economics*, **148**, 131–151.

Driscoll, John C., & Kraay, Aart C. 1994. Consistent covariance matrix estimation with spatially dependent panel data. *Review of Economics and Statistics*, **80**(4), 549–560.

European Union. 2018. Directive 2018/851/EC of the European Parliament and of the Council of 30 May amending Directive 2008/98/EC on waste. *Official Journal of the European Union 14/6/2018*, 109–140.

Fullerton, D., & Kinnaman, C. T. 1996. Household responses to pricing garbage by the bag. *The American Economic Review*, **86**(4), 971–984.

Fullerton, D., & Kinnaman, C. T. 2000. Garbage and recycling with endogenous local policy. *Journal of Urban Economics*, **48**(3), 419–442.

Grossmann, D., Hudson, J. F., & Marks, D.H. 1974. Waste generation models for solid waste collection. *Journal of the Environmental Engineering Division*, **100**(6), 1219–1230.

Hanley, R. 1992. Towns Adopt Pay-as-You-Throw Garbage. *The New York Times. Available at* *https://www.nytimes.com/1992/07/14/nyregion/towns-adopt-pay-as-you-throw-garbage.html*.

Hong, S., & Adams, R. M. 1999. Household responses to price incentives for recycling: Some further evidence. *Land Economics*, **75**(4), 505–14.

ISTAT. 2011. Population and Housing Census. *ISTAT - National Institute of Statistics. Data available at* `http://dati-censimentopopolazione.istat.it`.

ISTAT. 2018. Data on Environmental Crimes. *ISTAT - National Institute of Statistics. Data available at* `https://www.istat.it/it/archivio/218648`.

Jenkins, R. R., Martinez, S. A., Palmer, K., & Podolsky, M. J. 2003. The determinants of household recycling: A material-specific analysis of recycling program features and unit pricing. *Journal of Environmental Economics and Management*, **45**(2), 294–318.

Mentch, Lucas, & Hooker, Giles. 2016. Quantifying Uncertainty in Random Forests via Confidence Intervals and Hypothesis Tests. *Journal of Machine Learning Research*, **17**(26), 1–41.

Miranda, M. L., & Bauer, S. 1996. The urban performance of unit pricing: An analysis of variable rates for residential garbage collection in urban areas. *The U.S. Environmental Protection Agency*. Report prepared for the Office of Policy, Planning, and Evaluation.

Neyman, J. 1979. C(α) tests and their use. *Econometrica*, **41**, 1–21.

Pigou, A. C. 1932. *Divergences between marginal social net product and marginal private net product*. Vol. II, Chapter IX. London: Macmillan and Co.

Richardson, R. A., & Havlicek, J. H. 1978. Economic analysis of the composition of household solid wastes. *Journal of Environmental Economics and Management*, **5**(1), 103–111.

Robinson, P.M. 1988. Root-N-Consistent Semiparametric Regression. *Econometrica*, **56**(4), 931–954.

Rubin, Donald B. 1974. Estimating causal effects of treatments in randomized and non-randomized studies. *Journal of Educational Psychology*, **5**(5), 688–701.

Sexton, Joseph, & Laake, Petter. 2009. Standard errors for bagged and random forest estimators. *Computational Statistics and Data Analysis*, **53**(3), 801–811.

Tibshirani, Julie, Athey, Susan, Wager, Stefan, Friedberg, Rina, Miner, Luke, & Wright, Marvin. 2018. *grf: Generalized Random Forests (Beta)*. R package version 0.10.1.

Usui, T., & Takeuchi, K. 2014. Evaluating unit-based pricing of residential solid waste: A panel data analysis. *Environmental and Resource Economics*, **58**(2), 245–271.

Valente, M., & Bueno, M. 2018. The Effects of Pricing Waste Generation: A Synthetic Control Approach. *Under revision in the Journal of Environmental Economics and Management*.

Van Houtven, G. L., & Morris, G. E. 1999. Household behavior under alternative pay-as-you-throw systems for solid waste disposal. *Land Economics*, **75**(4), 515–537.

Wertz, K. L. 1976. Economic factors influencing households' production of refuse. *Journal of Environmental Economics and Management*, **2**(4), 263–272.