Regional Disparities in Waste Management: A Bayesian Analysis of Italian Municipalities

Gabriele Cabibbo 2196717

Abstract

This study investigates municipal waste management in over 4000 Italian municipalities trying to understand how consumers' waste habits are influenced by their income, different types of paying fees and different socioeconomic context in different parts of Italy. The analysis was conducted using data from the Bicocca Open Archive Research Data (BOARD) that includes geographical, economic, demographic, and waste management variables and applying Bayesian models by treating the geographical areas both as a sample of a broader, underlying population and as three distinct, unchangeable, and exhaustive divisions of Italy. The results from these Bayesian approaches were then compared with those obtained from frequentist methods, including a Multiple Linear Regression and a Linear Mixed-Effects Model. The results consistently show an heterogeneity in baseline waste per capita generation and that higher levels of income are associated with larger average individual outputs of waste. While Bayesian models showed a tendency for Pay-As-You-Throw (PAYT) schemes to decrease waste, the 95% credible intervals often included zero, suggesting no statistically significant effect. In contrast, estimating the effect of the adoption of a PAYT scheme using both frequentist approaches produced statistically significant negative effect. Overall, while regional differences and income effects are evident, the impact of PAYT schemes remains uncertain across different modeling approaches and a significant portion of waste per capita variability remains unexplained.

Contents

1 Introduction **2**

2 Data **2**

3 Exploratory Data Analysis **4**

4 Modelling **7**

4.1 Fixed Effect Model 7

4.2 Random Effect Model 10

5 Frequentist Approach **13**

5.1 Multiple Linear Regression 13

5.2 Linear Mixed-Effects Model 14

6 Conclusions **16**

7 References **18**

1 Introduction

Effective municipal waste management is a critical global challenge, encompassing environmental, economic, and social dimensions. As urban populations grow and consumption evolves, understanding the what drives solid waste generation becomes central for developing sustainable policies. This study focuses Italy, a nation characterized by significant and inherent socioeconomic differences between its northern, central, and southern regions and it investigates how different paying schemes and economic changes affect the per capita solid waste production, paying particular attention to the distinct socioeconomic contexts that are present in this nation.

I'll employ both Bayesian and Frequentist approaches to provide a comprehensive and robust analysis. Specifically, we utilize Bayesian **Fixed Effect** and **Random Effect** models to investigate how waste per capita is influenced by income and the adoption of Pay-As-You-Throw fee schemes, while accounting for inherent geographical variability.

The random effect model assumes that region-specific intercepts are drawn from a common distribution, allowing for the estimation of a population-level average and variability across regions. In contrast, the fixed effect model treats each regional intercept as an independent parameter.

To complement the Bayesian analysis, a **Frequentist Multiple Linear Regression** model is also employed, using dummy variables to explicitly account for geographical differences. Furthermore, a **Frequentist Linear Mixed-Effects** model is utilized to explicitly model the hierarchical structure of municipalities nested within geographical regions, allowing for the direct quantification of between-region variability and providing a frequentist counterpart to our Bayesian random effect approach.

2 Data

The dataset used was obtained from the Bicocca Open Archive Research Data (BOARD) and contains information about municipal waste management and related municipal revenues for over 4000 Italian municipalities. The dataset comprises 36 variables, covering geographical, economic, demographic, and waste management dimensions.

Geographical Variables

Spatial and territorial characteristics of each municipality:

- `region`: Name of the region
- `province`: Name of the province
- `name`: Name of the municipality
- `altitude`: Altitude above sea level (in meters)
- `isle`: Dummy variable indicating if the municipality is an island (1 = Yes, 0 = No)
- `sea`: Dummy variable indicating if the municipality is coastal (1 = Yes, 0 = No)
- `geo`: Geographical zone classification (1 = South, 2 = Center, 3 = North)

Economic Variables

These variables describe the financial and economic context of each municipality:

- `cost_pc`; Total waste management cost per capita
- `residual_cost_pc`: Residual waste management cost per capita

- `sorted_cost_pc`: Sorted waste management cost per capita
- `log_gdp`: Log of municipal GDP
- `log_income`: Log of taxable income in the municipality
- `payt`: Dummy variable indicating adoption of a "Pay-As-You-Throw" (PAYT) fee scheme (1 = Yes, 0 = No). This charge policy is a waste management system in which households are charged proportionally to the quantity of solid waste that they produce. Moreover, the PAYT scheme supplies incentives to recyclable waste collection.

Demographic and Infrastructure Variables

These variables provide demographic and urban context:

- `population`
- `area` (km^2)
- `population_density`
- `urb`: Urbanization level (1 = Low, 2 = Medium, 3 = High)
- `roads`: Total kilometers of road within the municipality
- `log_ppl_road`: Log of people per kilometer of road

Waste Composition and Management Variables

- `organic`, `paper`, `glass`, `wood`, `metal`, `plastic`, `electronics`, `texile`, `other`: Percentage of municipal solid waste by material type;
- `sorted_waste`: Sorted municipal solid waste (kg)
- `unsorted_waste`: Unsorted municipal solid waste (kg)
- `waste`: Total municipal solid waste generated (kg)
- `share_sorted_waste`: Share of sorted waste
- `share_w2e`: Share of solid waste sent to Waste-to-Energy (W2E) plants. These plants perform a series of processes designed to convert waste materials into usable forms of energy. The above facilities represent a pivotal role in waste management and sustainable energy production since they both reduce the volume of litter in landfills and provide alternative energy sources.
- `share_landfill`: Share of waste sent to landfill

With four other administrative and coding variables which won't be used in the following analysis.

Table 1: Number of Missing Values per Variable

Variable	Missing Values	Percentage
residual_cost_pc	52	1.20
sorted_cost_pc	67	1.54
area	6	0.14
altitude	6	0.14
isle	6	0.14
sea	6	0.14
pop_density	6	0.14
waste_per_sq_km	6	0.14
urb	6	0.14
organic	512	11.79
paper	25	0.58
glass	33	0.76
wood	1095	25.22
metal	246	5.67
plastic	39	0.90
electronics	314	7.23
texile	1013	23.34
other	136	3.13
geo	285	6.57
roads	443	10.21
share_w2e	285	6.57
s_landfill	285	6.57
log_gdp	386	8.89
log_ppl_road	443	10.21
log_income	285	6.57
finance	386	8.89

3 Exploratory Data Analysis

Table 1 reveals significant missing data across many columns. Dropping these incomplete records would result in a large loss of information, leaving only 2017 observations.

To avoid this substantial loss of data, continuous variables (excluding pre-existing logarithmic scales and shares) were log-transformed before being imputed in a multivariate way using the `mice` library. Share variables have been imputed in the same way but were first transformed in logit values rather than logarithms.

Unlike continuous variables, categorical variables presented fewer missing entries. The main exception is the `geo` variable which had a notable number of missing values. To account for these, the covariate has been imputed by determining the region in which each observation is situated. Additionally, in this variable took some illogical values for some units, i.e. some values were equal to 1.5, since these were located either in Marche or in Tuscany, they were re-labelled as "2" (Center).

In this way we're left with just 6 units containing missing values and we can drop them without losing too much information.

After imputing missing data, I defined waste per capita for each city that will be used as a response variable for subsequent modelling.

The distributions of waste per capita across Italy's three macro-regions (Figure 1) appear similar. However, the boxplot reveals that Southern cities, on average, generate less waste per capita compared to their Northern counterparts. All the macro-regions share a large number of extremely high waste per capita values, with 2

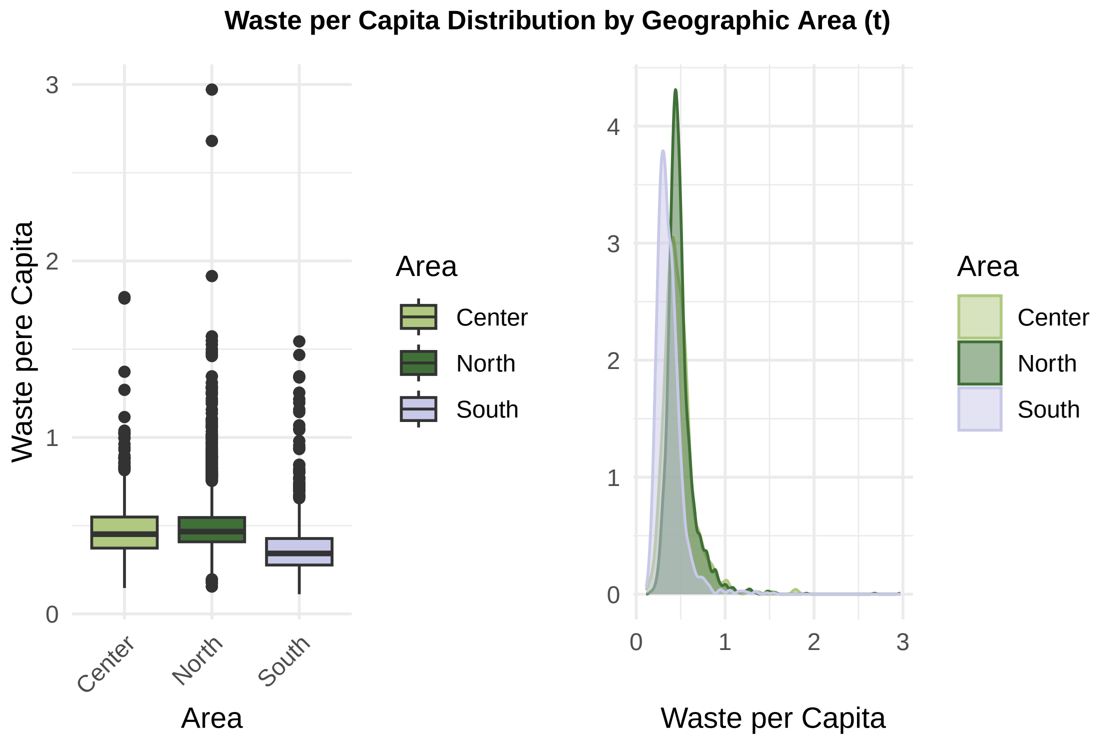

Figure 1: Waste per capita Distribution by Geographical Area

Northern cities standing out: Limone sul Garda and Lignano Sabbiadoro, this is also the only geographical area showing two outliers taking remarkably small values which are represented by the cities of Lauregno and San Pancrazio, both in Trentino Alto Adige. Nevertheless, we can notice that even if these municipalities stand out for their low waste per capita production, they're not the cities that assume the smallest value for this variable, which are instead Carpineto della Nora and Campora in Abruzzo and Campania, respectively. In absolute terms, Rome produces the most waste and Maccastorna (Lombardia) the least, but they do not stand out for their per capita waste.

Figure 2 illustrates the average per capita waste production by administrative region. As previously observed, Northern cities generally show higher average solid waste output per citizen, while Southern regions produce less. Central regions exhibit more intermediate levels, with the exclusion of Tuscany, which is the second highest region with almost 600 kg of waste per person. This distribution is in contrast with the absolute average solid waste produced per Italian region, where we cannot distinguish any North-South-Center pattern.

Figure 3 presents the distribution of the logarithm of the municipalities' income. Surprisingly, Southern cities, on average, display a higher log income. This observation can be attributed to the prevalence of very small towns in Northern Italy, whose lower incomes pull down the empirical average for that geographical area. Nevertheless, the income distribution for Southern cities shows two extremely low values, while the Center and Northern distributions exhibit a large number of extremely large values.

Finally, comparing waste per capita in cities with a standard paying fee versus those adopting a Pay-As-You-Throw policy (Figure 4), we can't notice notice any major difference. Both density plots peak around 500 kg per person and exhibit long, right tails. However, municipalities with standard paying schemes tend to be more skewed towards lower waste values. We also notice that the adoption of the PAYT scheme is extremely unbalanced: almost 94% of the cities that adopt this policy are in Northern Italy, while less than 1% of the total adoptions are in Southern cities.

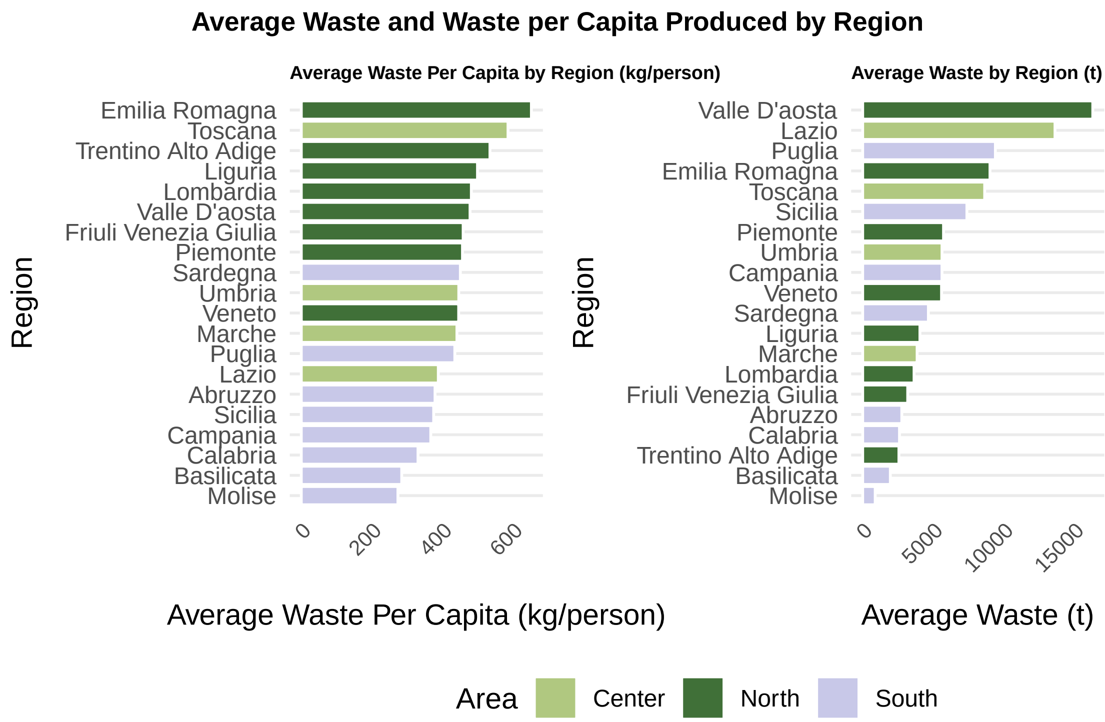

Figure 2: Average Waste and Waste per Capita Produced by Region

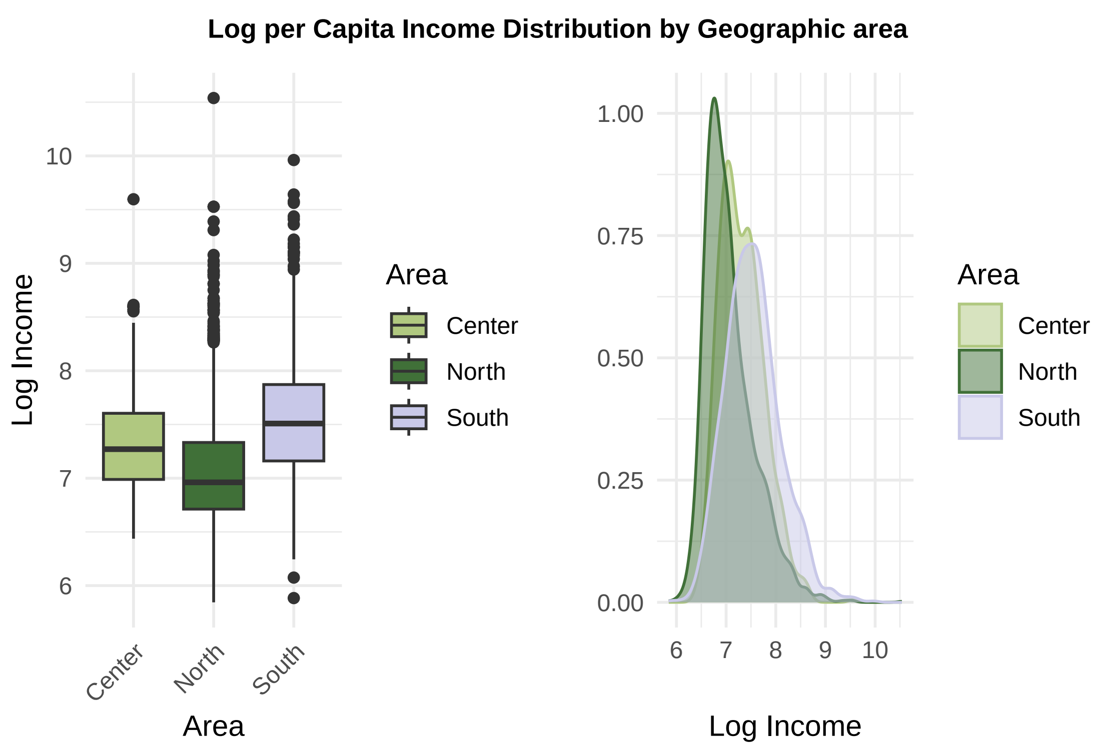

Figure 3: Log Per Capita Income Distribution by Geographical Area

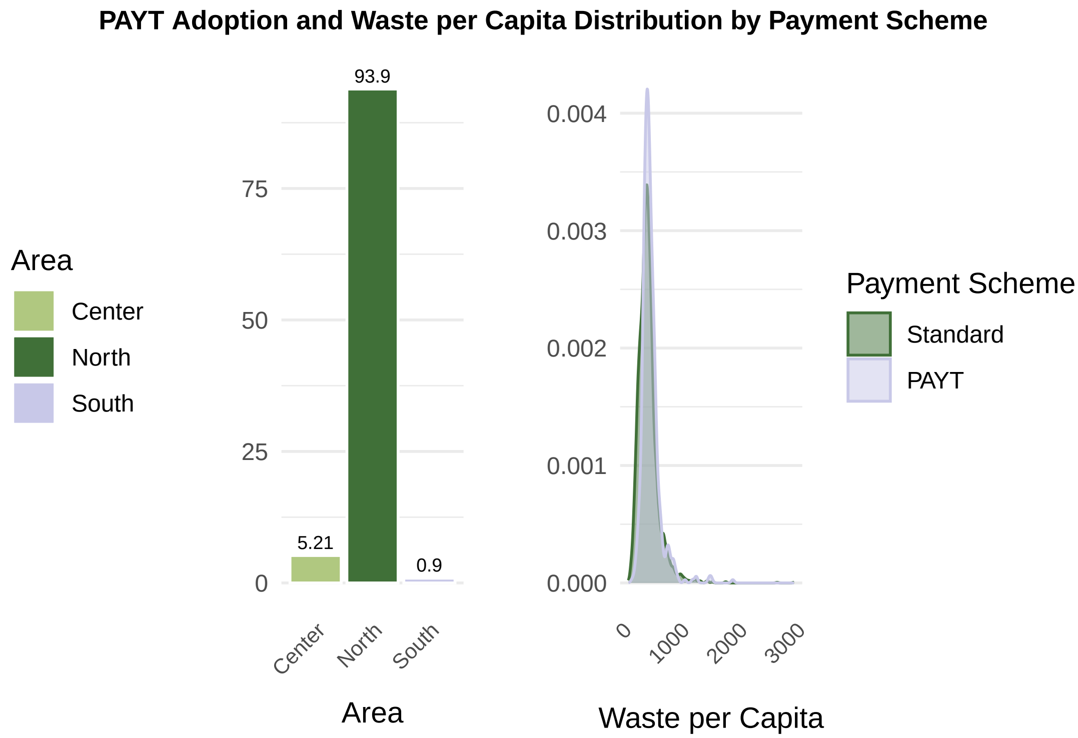

Figure 4: PAYT Adoption and Waste per Capita Distribution by Payment Scheme

4 Modelling

In this section I'm going to investigate how per capita waste production is affected by the (log transformed) per capita municipal income and the introduction of a Pay-As-You-Throw fee scheme, while critically accounting for the inherent geographical variability across Italian municipalities.

I will explore two different but related Bayesian frameworks: **Fixed Effect Models** and **Random Effect Models**. The former approach treats each level of a categorical grouping variable (the geographical areas) as distinct and independent entities, estimating a separate, unique intercept for each group, effectively modeling the specific average waste per capita for that particular region assuming all other covariates are constant. The latter one, instead, views the groups as a sample drawn from a larger population of possible groups. It assumes that these group-specific intercepts themselves vary according to a common underlying probability distribution and allows groups with limited data to "borrow strength" from those groups which are more data rich.

4.1 Fixed Effect Model

4.1.1 Model Setup

This model aims to explain how the solid waste produced by each citizen changes with income and how it varies in places in which a Pay-As-You-Throw policy is adopted. Here each geographical area is treated as fixed and distinct and the intercept relative to each area is estimated independently.

In this framework each observation i is distributed accordingly to the following model:

$$Y_i \sim Normal(\mu_i, \tau_Y)$$

where μ_i is defined as:

$$\mu_i = \beta_{0,\ geo[i]} + \beta_{payt} X_{payt,i} + \beta_{income} X_{income,i}$$

with $\beta_{0,\ geo[i]}$ is a fixed intercept for the specific geographical area in which the observation is situated and each of them has an independent, weakly informative, prior distribution with the following form: $Normal(0, 0.001)$. A prior of the same form is analogously assigned for β_{payt} and β_{income}.

Ultimately, τ_Y is the precision of the residuals and is computed starting from the standard deviation, whose prior is a weakly informative uniform prior: $\sigma_Y \sim Uniform(0, 500)$.

4.1.2 Model Results

Each geographical area gets its own, independent intercept β_{geo}, each one representing the average waste produced per capita when the logarithm of income is centered at its mean and a city adopts a standard fee scheme.

Table 2: Fixed Effect Model Summary

Parameter	**mean**	**sd**	**2.5%**	**50%**	**97.5%**	**Rhat**	**n.eff**
$\beta_{0,\ South}$	422.708	5.918	411.201	422.712	434.255	1.001	24000
$\beta_{0,\ Center}$	462.467	6.538	449.706	462.489	475.271	1.001	24000
$\beta_{0,\ North}$	464.784	4.178	456.586	464.788	473.085	1.001	24000
β_{income}	179.225	11.840	156.159	179.159	202.558	1.001	24000
β_{payt}	-6.517	7.451	-20.969	-6.506	7.994	1.001	23000
deviance	56387.557	10.401	56369.993	56386.656	56410.568	1.001	24000
σ_Y	161.580	1.741	158.191	161.567	165.064	1.001	18000
DIC: 56441.6							

The data shows distinct baseline per capita waste levels across different parts of Italy, displaying that cities of northern Italy are observed to have the highest waste outputs, even if the estimate is just 2 kg per capita over the cities in central Italy, while southern people show a smallest baseline waste production.

Moving our focus to common coefficients, we notice that the estimate for the mean of β_{income} close to 180, implying that a 1% increase in the average citizen's income is associated with an increase close to 1.8 kg in their annual waste production. We also observe a negative mean estimate for β_{payt}, but the 95% credible interval contains both positive and negative values, hence we can't conclude that there is a strong evidence for a statistically significant negative effect of the Pay-As-You-Throw scheme on the citizens' waste habits.

Finally, σ_Y measures the standard deviation of the residuals, reflecting typical amount of waste per capita variability that isn't explained by either the Pay-As-You-Throw scheme or the logarithm of income. The large mean that we observe points to considerable unexplained variability within regions, even after controlling for the included covariates. This suggests that other factors that have not been included in this model, are also influencing waste generation.

4.1.3 MCMC Convergence Diagnostics

To ensure the reliability of our posterior inferences we have to inspect the Monte Carlo Markov Chain diagnostics.

For each parameter, the potential scale reduction factor ($\hat{R}$) is equal 1.001 (for proper convergence, $\hat{R}$ values should be very close to 1.0), showing robust sampling and reliable parameter estimates and providing strong evidence for the convergence of all MCMC chains. This confirms the stability of our estimates and suggests that the sampling process adequately explored the parameter space.

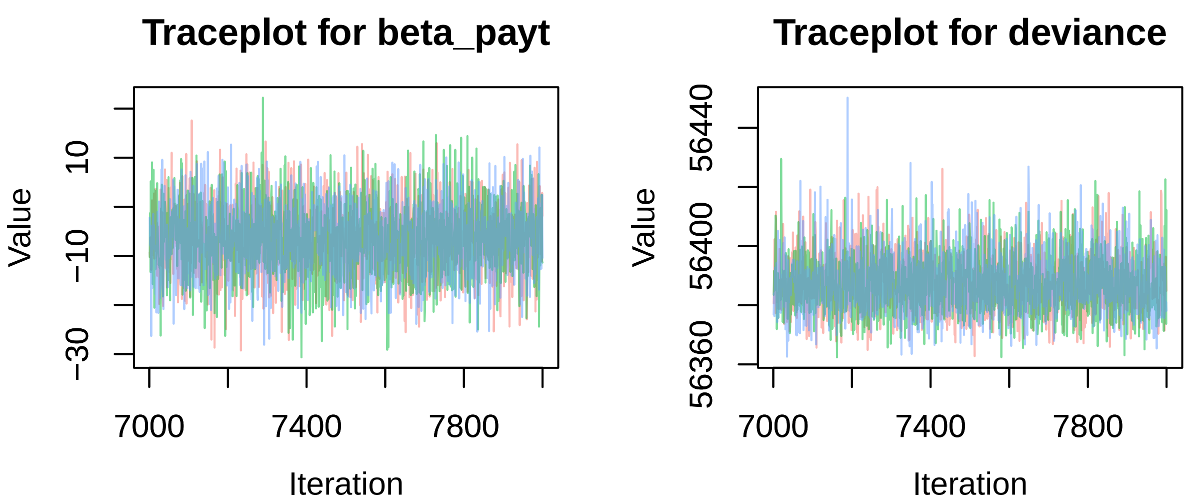

Figure 5: Traceplots of β_{payt}, and deviance for the final 1000 iterations of the MCMC sampling.

Most of the effective sample sizes obtained are equal to 24000, indicating low autocorrelation and highly efficient sampling. The effective sample size for β_{payt} is only slightly smaller, while σ_Y has an n_{eff} of 18000, which is still sufficiently large in absolute terms, ensuring precise and reliable estimates for its posterior distribution.

Visual inspection of the traceplots (Figure 5) shows good mixing of the chains, regularly displaying the desired "hairy caterpillar" appearance. However, we have to notice that the traceplots for β_{payt} and, more notably, for *deviance* displayed some occasional extreme values.

4.2 Random Effect Model

4.2.1 Model Setup

In this section, I develop a model to understand the determinants of per capita solid waste. My focus remains on the impact of income and Pay-As-You-Throw policies. This time, however, the geographical areas are assumed to be drawn from a larger population of possible groups which allows us to "borrow strength" across regions and estimate the overall extent of geographical variability in waste generation. This is an important property that we want to exploit, particularly given the extreme imbalance previously observed in PAYT adoption across different geographical areas.

For each observation in the j-th geographical area the solid waste produced per capita is distributed as:

$$Y_i \sim Normal(\mu_i, \tau_Y)$$

in which μ_i is defined as:

$$\mu_i = \beta_{0,\ geo[i]} + \beta_{payt} X_{payt,i} + \beta_{income} X_{income,i}$$

with $\beta_{0,\ geo[i]}$ being the region-specific intercept, the other two β parameters represent the common coefficients for `payt` and `log_income`.

The three region-specific intercepts are assumed to be drawn from the same normal distribution parametrized by a mean μ_{β_0} and precision τ_{β_0} in which the former has a weakly informative prior distribution $\mu_{\beta_0} \sim Normal(0, 0.001)$ (which is also the prior for the β parameters for `payt` and `log_income`) and the latter is computed starting from the prior distribution of the standard deviation which is once again weakly informative Half-Normal distribution that ensures positivity: $\sigma_{\beta_0} \sim Normal(0, 0.0001)$.

Finally, τ_Y is the precision of the residuals computed from their standard deviation σ_Y which is distributed as a weakly informative $Uniform(0, 500)$.

4.2.2 Model Results

Table 3 shows the estimated parameters of the random effect model for waste per capita.

Each $\beta_{0,\ geo}$ represents the average waste per capita the cities in the three distinct Italian geographical areas. These averages are estimated when a standard paying fee is adopted (`payt=0`) and the logarithm of per capita income is at its mean. The results clearly show different baseline waste production across these three areas of Italy. In contrast with previous results, citizens in central regions appear to produce the most waste on average, while those in the South produce the least. This highlights intrinsic regional differences in waste generation patterns.

The β_{income} coefficient quantifies how changes in income affect an individual's average waste output. The estimate for this coefficient is equal to 188.53, suggesting that a 1% increase average citizen's income is associated with an increase of almost 2 kg in their annual waste production. The magnitude of this coefficient, measures a slightly larger impact of changes in income on waste per capita output.

Table 3: Random Effect Model Summary

Parameter	**mean**	**sd**	**2.5%**	**50%**	**97.5%**	**Rhat**	**n.eff**
$\beta_{0,\ South}$	434.124	6.025	422.178	434.118	445.931	1.001	24000
$\beta_{0,\ Center}$	483.208	6.683	470.150	483.192	496.299	1.001	24000
$\beta_{0,\ North}$	469.513	4.170	461.301	469.510	477.683	1.001	24000
β_{income}	188.531	11.974	165.096	188.571	211.994	1.001	11000
β_{payt}	-13.738	7.447	-28.294	-13.734	0.900	1.001	24000
deviance	56367.063	5.684	56358.370	56366.274	56380.292	1.001	6800
$\mu_{\beta 0}$	20.570	31.961	-42.215	20.558	83.126	1.001	14000
σ_Y	161.179	1.725	157.861	161.156	164.616	1.001	24000
$\sigma_{\beta 0}$	263.820	48.065	182.304	259.386	369.832	1.001	20000
DIC: 56383.2							

The β_{payt} coefficient aimed to explain the average effect of introducing a Pay-As-You-Throw (PAYT) scheme. This estimate suggests a decrease of almost 14 kg in an individual's waste habits. However we notice that the 95% credible interval for β_{payt} is $(-28.29, 0.9)$, which contains zero. This means that while there is a tendency for this fee scheme to be associated with lower personal waste, our data does not provide strong evidence to rule out a null effect or even a small positive effect.

Moving to the population-level hyperparameter estimates, we can notice that μ_{β_0}, the estimated mean of the region-specific intercepts across all possible regions, shows considerable uncertainty. Its very wide 95% credible interval comprises both negative and positive values, indicating a considerable uncertainty about the overall average baseline across all potential regions. In contrast, the estimate for the standard deviation of the region-specific intercepts, σ_{β_0}, presents a large 95% credible interval containing many high magnitude values, strongly supporting the presence of a substantial heterogeneity in baseline waste generation across regions. This confirms that regional differences in waste habits are a notable factor.

4.2.3 MCMC Convergence Diagnostics

The Monte Carlo Markov Chain diagnostics relative to the random effect model show robust sampling and reliable parameter estimates. As we can observe from Table 3, the potential scale reduction factor for each parameter is 1.001, providing strong evidence that all the MCMC chains have converged to their target posterior distribution. This indicates that the parameter space has been properly explored and that the posterior estimates are stable and trustworthy.

For most parameters we can notice an effective sample size (n_{eff}), which quantifies the number of independent samples obtained from the posterior distribution, larger than 20000, this large value means that the sampling has been efficient and that the parameter estimates are precise.

For some hyperparameters, such as β_{income}, μ_{beta_0} and *deviance*, the effective sample sizes recorded are slightly lower but are still substantially large in absolute terms, indicating precise and reliable estimates of the posterior distributions for these parameters. Figure 6 shows the traceplots relative to the final 1000 iterations of the sampling process for these hyperparameters, we can notice that they present the typical hairy caterpillar shape, indicating of good mixing. However, the presence of occasional extreme values might explain their comparatively lower n_{eff} values.

In conclusion, the MCMC diagnostics confirm that the model's sampling procedure was highly effective, yielding converged chains and a large number of independent samples. This assures us that the subsequent interpretations of the model parameters are based on robust and reliable posterior estimates.

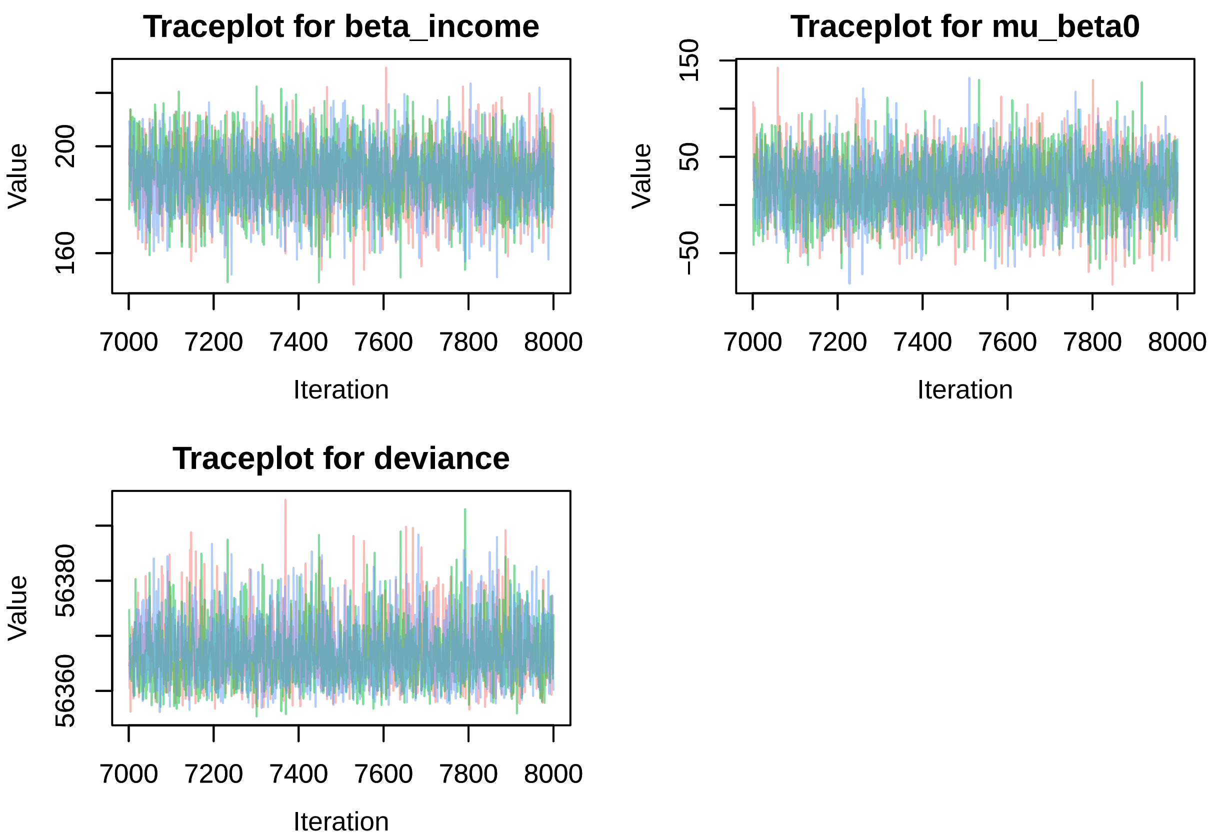

Figure 6: Traceplots of μ_{β_0}, β_{income}, and deviance for the final 1000 iterations of the MCMC sampling.

5 Frequentist Approach

5.1 Multiple Linear Regression

5.1.1 Model Setup

Now we switch to a frequentist approach that aligns with the structure of the previous Bayesian fixed effect model, a multiple linear regression which uses dummy variables to define different intercepts for each geographical area, as represented by the equation:

$$Y_i = \beta_0 + \delta_{North}\beta_{North} + \delta_{South}\beta_{South} + \beta_{payt}X_{payt,i} + \beta_{income}X_{income,i} + \epsilon_i$$

In which:

- Y_i is the waste per capita produced in the i-th municipality
- β_0 is the intercept for the reference geographical area (Center)
- δ_{North} and δ_{South} are the dummy variables representing Northern and Southern cities, respectively
- β_{North} and β_{South} are the coefficients representing the difference in intercepts for North and South regions compared to the Central one
- $\beta_{payt}X_{payt,i}$ is the coefficient that measures the difference in the averege waste per capita in cities which adopt a Pay-As-You-Throw fee scheme with respect to those who apply a standard policy
- β_{income} is the coefficient relative to the logarithm of the income
- ϵ_i is the error term.

5.1.2 Results

Table 4 shows the estimated coefficients of my model: the intercept, β_0 measures the baseline average waste produced by each citizen in a municipality belonging to the reference geographical area (Center) under a standard fee scheme and the log_income is at its mean. Its estimated value is close to 484 kg per capita.

The difference in average per capita waste output for the Northern and Southern cities compared to the ones in Italy's center has been estimated. Both this coefficient estimates result negative, indicating less waste per capita for municipalities in the North and South compared to the Center, keeping all other factors constant.

β_{payt} measures the estimated difference in waste per capita between cities which adopt a Pay-As-You-Throw policy in opposition to those which apply a standard paying scheme. We can observe that this coefficient is negative, showing that the adoption of a PAYT scheme is associated with a reduction in citizens' waste habits. Keeping all other factors constant, the inhabitants of cities with PAYT policies on average produce almost 16.5 kg less of solid waste per capita.

Finally, we estimated a β_{income} coefficient equal to 220, meaning that a one percent increase in income is expected to increase the waste per capita by slightly more than 2 kg.

We can also notice that all the coefficients are statistically significant at a 5% level, meaning that we have strong evidence that our explanatory variables have a non-zero effect on the waste per capita. However an adjusted R^2 equal to 16.76% shows that our model explains only a small fraction of the total variablity of our response variable.

Table 4: Frequentist Multiple Regression Results

	Dependent variable:
	Waste per Capita
Intercept	483.902***
	(6.693)
Area: North	−19.632**
	(7.893)
Area: South	−38.860***
	(9.100)
Pay-As-You-Throw	−16.432**
	(7.684)
Income	220.316***
	(12.894)
Adjusted R^2	0.168
Residual Std. Error	161.041 (df = 4330)
F Statistic	219.181*** (df = 4; 4330)
Note:	*p<0.1; **p<0.05; ***p<0.01

5.2 Linear Mixed-Effects Model

5.2.1 Model Setup

Finally, we estimate a frequentist Linear Mixed-Effect Model which includes both fixed and random effects. In this framework, we compute an overall fixed intercept that measures the average baseline waste per capita across all geographical areas when the logarithm of income assumes its mean value and the city adopts a standard paying scheme, together with a random intercept for each geographical area that quantifies the deviation of the geographical area's baseline waste per capita from the previous intercept.

This model assumes the following form:

$$Y_i = \left(\beta_0 + u_{0,j}\right) + \beta_{income} X_{income,i} + \beta_{payt} X_{payt,i} + \epsilon_{i,j}$$

Where:

- Y_i is the waste per capita produced in the i-th municipality
- β_0 is the overall fixed intercept
- $u_{i,j}$ is the random intercept for the j-th geographical area
- β_{income} and β_{payt} are the fixed effect coefficients relative to the logarithm of income and the adoption of the PAYT scheme, respectively
- $\epsilon_{i,j}$ for the residual error for the i-th municipality in the j-th geographical area.

5.2.2 Results

Table 5 shows the results for my Linear Mixed-Effects Model. The left panel of the table shows the fixed effects part of the model: the estimate for the overall fixed intercept is close to 465 kg of waste per person, which is in line with the results we obtained in the previous models. The coefficient estimates for β_{payt} and β_{income} are also very similar to the previous estimates. We notice that, as in the prior frequentist model, the estimate of the effect of a Pay-As-You-Throw policy on the per capita waste production is negative and statistically significant at a 5% level.

The random effects estimates are displayed in the right panel of the table: the intercept relative to Northern cities is very close to the overall mean, while the intercepts for the other two areas are much further away, showing that the average baseline waste per capita production in Central Italy is way larger than the overall baseline and a lot bigger than the one observed for Southern cities.

Table 5: Frequentist Linear Mixed-Effects Model Results

	Fixed Effects	**Random Effects**
Intercept	464.114*** (10.942)	Center: 17.453
Income	223.139*** (12.497)	North: −0.350
Pay-As-You-Throw	−16.457** (7.676)	South: −17.103
		Std. Dev. (Intercept): 18.22
		Std. Dev. (Residual): 161.04
Log Likelihood: −28173.140		
AIC: 56356.290		
BIC: 56388.160		

Note: *p<0.1; **p<0.05; ***p<0.01

6 Conclusions

This analysis provides valuable insights about policy and economic related factors affecting waste generation in the complex economic Italian landscape. The exploratory data analysis that I conducted highlights mild regional disparities described by a generally higher average per capita waste production for Northern cities compared to their Southern counterparts. This analysis also depicts a substantial uncertainty of the effects of Pay-As-You-Throw policies on the average citizen waste habits.

Table 6: Estimated Baseline Waste Per Capita by Geographical Area Across Models

Area	**Bayesian**	**Frequentist**
Fixed Effects Models		
South	422.708	445.082
Center	462.467	483.902
North	464.784	464.270
Random Effects Models		
South	434.124	447.011
Center	483.208	481.567
North	469.513	463.764

The **Bayesian Random Effect** model captures an underlying heterogeneity in baseline waste per capita generation across Italy's different geographical area, suggesting inherently different waste habits associated with distinct socioeconomic contexts. The coefficient measuring the relationship between the logarithm of income and the waste per capita production demonstrated a clear positive influence of the covariate on our response variable. This displays the impact of economic prosperity on people's consumption habits and, thus, waste output. The model also shows a clear tendency for the PAYT fee scheme to be associated with a smaller waste per capita production, but the data analyzed does not provide strong evidence to rule out a null or even a small positive effect.

Our **Bayesian Fixed Effect** model, shows similar results, highlighting distinct baseline levels in waste per capita in the three geographical area and, again, an higher waste per capita outputs that are associated with higher income levels. This model produces, as well, uncertain results for the effect of the Pay-As-You-Throw fee scheme on each citizen's average waste habits.

Implementing a **frequentist** approach, I obtained some results that match our Bayesian approaches and others that are fairly in contrast with them. A **Multiple Linear Regression** model shows, once again, different baseline waste production in distinct geographical areas and an increase in waste per capita in cities presenting higher incomes levels, keeping all other factors fixed. The main difference between our approaches is that, estimating the coefficients using frquentist methods, we obtain a β_{payt} that is both negative and statistically significant (at a 5% level), suggesting smaller individual waste production in municipalities adopting a PAYT paying scheme.

Finally, a **Frequentist Linear Mixed-Effects Model** largely supports the insights obtained from the other models: effectively quantifying the significant between-region variability in baseline waste production, confirming the presence of distinct regional waste habits. Like the other models, it robustly indicates that higher income levels are associated with increased waste per capita and, as in the previous frequentist approach finds a statistically significant association between the adoption of a PAYT fee scheme and a decrease in waste per capita levels (at a 5% level).

A consistent finding across all models was a substantial unexplained variability in waste per capita, indicated by the large standard deviation of the residuals σ_Y. This suggests that while geographical area, income, and the adoption of PAYT schemes are relevant factors, they do not fully account for the observed differences in waste generation and that many other unobserved factors likely play significant roles. This is also reinforced by the R^2 obtained in the frequentist context, which tells us that our model explains 16.76% of the total variabilty of the response variable.

We can, thus, remark the importance of different socioeconomic contexts in waste management strategies and the positive correlation between income and waste generation, while alternative paying schemes, like a Pay-As-You-Throw, show promise in encouraging waste reduction but their effectiveness is subtle.

7 References

Foggia, G.D. (2022) Municipal Waste Management Cost and Fee Schemes, Mendeley Data. Available at: https://board.unimib.it/datasets/w5f9kg7743

Gelman, A et al. (2013) Bayesian data analysis. 3rd edn.

https://en.wikipedia.org/wiki/Waste-to-energy

https://en.wikipedia.org/wiki/Pay_as_you_throw

https://greenbestpractice.jrc.ec.europa.eu/node/7