Author Accepted Manuscript version of the paper by Tuva Grytli and Cansu Birgen
in Computer-aided chemical engineering, Vol 52 (2023) Page 2149-2154,
DOI: http://dx.doi.org/10.1016/B978-0-443-15274-0.50342-5

A Critical Look at Waste Composition Analyses: Challenges and Opportunities

Tuva Grytli,[a] Cansu Birgen,[b]

[a] *SINTEF Industry, P.O. Box 4760 Torgarden, Trondheim, 7465, Norway*
[b] *SINTEF Energy Research, P.O. Box 4760 Torgarden, 7465 Trondheim, Norway*

Abstract

This paper presents regression analysis results on time series data of waste composition analyses from a municipal waste company in Norway, with a discussion on methodological challenges. The aim is to investigate what affects municipal solid waste (MSW) source-sorting results, to facilitate achieving circular economy targets. The share of four MSW fractions in the mixed waste bin are studied considering five independent variables: Collection/sorting systems for food waste, glass and metal packaging waste, and mixed waste; year of analysis, and pre/post covid-19. We find that only the source sorting systems have significant impacts on the waste composition, and that user-friendliness matters. We recommend more in-depth studies to increase data quality to better utilise the potential of waste composition analyses for a more circular waste system.

Keywords: Municipal solid waste, sorting behaviour, waste management, regression analysis, circular economy.

1. Introduction

To improve circularity, recycling, and reuse, we need to improve municipal solid waste (MSW) management systems, where large quantities of materials have the potential to be recovered. The 2021 material recovery rate in Norway was 43 %, well below the 55 % target for 2025 (Statistics Norway 2020, Directive (EU) 2018/851 2018). More than 60 % of the waste in the mixed waste (MW) bin is recyclable, so understanding how households sort their waste is vital, and good data is key for effective waste strategy planning (Grytli & Birgen 2023). The aim of this study is to increase the insight into what affects sorting behaviour, to enable increased circularity in the waste treatment system.

This study is based on waste composition analysis data from a municipal waste company in Norway that serves more than 100 000 inhabitants. Waste composition analyses of the MW bin were carried out in 2016, 2017, 2019 and 2021, comprising a time series where in total almost 4 000 kg waste has been physically analysed.

2. Methodology

To better understand what affects waste sorting at source, the results from the waste composition analyses were investigated using multiple linear regression analysis. The number of sample areas analysed per year were two (2016-2017), three (2019) and four (2021), giving a total of 11 observations.

The waste composition analyses examine the contents of the MW bin, estimating the share, in weight percentage, of ten different waste fractions. Table 1 shows the sample area average share per fraction. A regression analysis was run for each fraction, with the share of the fraction in the bin as the dependent variable. Significant effects were found for five fractions, and this study focuses on four of them: Food waste (FW), glass packaging waste (GW), metal packaging waste (MeW), and MW which is the correctly sorted MW.

Table 2 Overview and average share of waste fractions in the mixed waste bin.

Waste fractions with average shares	
Paper and cardboard	7.65 %
Food waste incl. paper towel	43.6 %
Garden waste	4.83 %
Plastic packaging	9.43 %
Glass packaging	2.94 %
Metal packaging	1.43 %
Other metal	0.98 %
Recyclable textiles	3.03 %
Hazardous waste and WEEE	1.35 %
Mixed waste	24.8 %

To identify independent variables, the differences between the 11 observations were assessed. Two sample areas have separate FW collection. Kerbside glass and metal packaging waste (G&M) collection was rolled out between 2017 and 2019, and five observations took place after this roll-out. Three sample areas have underground MW collection, where the waste is discarded in a large underground container via a waste inlet aboveground. The other eight have regular rolling bins for MW. Four observations were from late fall of 2021, i.e. after Covid-19. These differences gave rise to five potential independent variables, described in Table 2. As all independent variables (except year) are categorical, they were coded as dummy variables (Wooldridge 2009).

Table 1 Overview and coding of independent variables.

Independent variables	*Short name*	Values			
		0	1		
Separate FW collection	*sepFW*	No separate collection	Separate FW bin		
Kerbside G&M collection	*kerbG&M*	G&M to central recycling point or civic amenity site	Kerbside G&M collection		
Underground waste collection	*Undergr.*	Aboveground MW bin	Underground MW system		
Post Covid-19	*Covid19*	Before Covid-19	After Covid-19		
		1	2	4	6
Year	*Year*	2016	2017	2019	2021

There were some challenges with the identified independent variables. Most had some correlation with the time variable, as they either happened at a certain time (kerbside G&M collection and Covid-19) or were included only in later years (underground waste collection). This can cause a multicollinearity problem, as discussed later. The next challenge was a very small number of observations for a regression analysis. Each observation is however the aggregated waste from 30-40 bins. This means that the 11 sample areas represent more than 330 bins, and almost 400 households. For practical reasons this study was done on a sample area level. It is unlikely that additional resolution would have added much value, as we had no information about each household, thus could not have explained any additional variation. As each bin is a snapshot in time and may not be representative of the household’s average waste composition, an average of a sample area may be considered more representative of an average waste composition. In general, a small number of observations increases uncertainty and means that the results should be interpreted carefully and may not be generalisable.

During analysis, unexpected results emerged. Close inspection of the data revealed large increases in one fraction, that could not be explained even after communication with the

MSW company. Thorough debugging revealed a summation error for a subfraction of the fraction in a spreadsheet, causing a hard to spot error that had significant impacts on the results. Data from waste composition analyses is registered on paper and punched manually into a spreadsheet template. As a copy of the template was used, the mistake spread to five sample areas over two years, causing a systematic error. Left undiscovered, this systematic error would have impacted the accuracy of the model significantly. Several random errors were also discovered, where numbers were punched incorrectly. Although most random errors were too small to affect overall results, they would have reduced the model's precision.

Due to Norway's sparse and spread population and complex geography, MSW companies are often small, which induces several of the methodological complications discussed. Small companies have limited resources for carrying out waste composition analyses, resulting in fewer observations, and for quality assuring results. The uncovered systematic error gave a faulty picture of the company's performance in collection, sorting and recycling, ultimately giving an incorrect base for their strategic development. This shows the benefit provided by research, as in addition to giving insight into effects of measures implemented, the MSW company's data is improved and quality assured.

3. Results and discussion

Table 3 Results for food waste fraction regression model.

Dependent variable: **Food waste**					
R^2	91 %	Adj. R^2	82 %	*95 % confidence intervals*	
Obs.	11	*Coeff.*	*P-value*		
(Intercept)		53 %	< 0.01	43 %	63 %
sepFW		-25 %	< 0.01	-40 %	-10 %
kerbG&M		7.4 %	0.37	-12 %	27 %
Undergr.		0.3 %	0.95	-10 %	11 %
Covid19		-6.7 %	0.40	-25 %	12 %
Year		-1.5 %	0.40	-6 %	3 %

The results from the linear regression model for the FW fraction are shown in Table 3. The model explains (R^2) 91 % of the variation in the fraction, but the only variable with a significant effect ($p<0.01$) is separate FW collection. The model finds that the share of FW is reduced by 25 percentage points (pp) when there is separate FW collection, from 53 % to 28 % FW in the MW bin, amounting to a 48 % reduction. The confidence intervals indicate that there is a 95 % likelihood that the reduction is between 10 and 40 pp. None of the other variables were found to have a significant effect on the share of FW in the MW bin.

Multicollinearity (a correlation between several of the independent variables) may give unreliable results in a regression analysis. To assess whether multicollinearity was a problem in the model, the variance inflation factors (VIF) for the independent variables were analysed (Wooldridge 2009). There is no agreed-upon threshold for when multicollinearity becomes problematic, but threshold values of 5 or 10 are commonly used. The VIFs for the independent variables are never over 10, but for kerbside G&M collection, year and Covid-19, the VIF is over 5. This became a problem for the GW and MeW fractions, where results were significantly affected. The collection system for G&M has been shown to impact the sorting behaviour in other studies (Grytli & Birgen 2023, Syversen et al 2019), and it is therefore assumed to be more important than development over time, or the effect of a pandemic. To avoid the multicollinearity problem, the variables Post Covid-19 and Year were excluded from the models for the GW and MeW fractions, and a simplified model is presented.

The results for the simplified regression models are shown in Table 4 for GW and Table 5 for MeW. For GW we found that the adjusted R^2 increased when running the model with only kerbside G&M collection as independent variable, indicating that the simplified model may be overfitted. Results from both runs are thus included in Table 4, and the discussion will focus on the *Only G&M* model for GW. The underlying reasons will be discussed later.

Table 5 Results for glass packaging waste fraction regression models.

Dep. variable: **Glass pack. waste**		Simplified model				Only G&M			
		R^2	54 %	Adj. R^2	34 %	R^2	50 %	Adj. R^2	45 %
Obs.	11	*Coeff.*	*P*	*95 % CI*		*Coeff.*	*P*	*95 % CI*	
(Intercept)		3.7 %	< 0.01	2.6 %	4.8 %	3.7 %	< 0.01	2.8 %	4.7 %
sepFW		0.8 %	0.51	-1.9 %	3.5 %				
kerbG&M		-2.2 %	0.04	-4.3 %	-0.1 %	-1.8 %	0.01	-3.1 %	-0.4 %
Undergr.		0.4 %	0.65	-1.6 %	2.5 %				

For GW the model explains (R^2) 50 % of the variation in the fraction, while the MeW model explains more of the variation, with an R^2 of 66 %. Only the kerbside G&M collection affects the share of G&M waste in the MW bin significantly in any of the models. For MeW the effect is more clearly significant ($p<0.01$) than for GW ($p=0.01$). The estimated reduction in MeW is 0.9 pp, from 1.6 % to 0.7 % MeW in the MW bin, amounting to a 55 % reduction. The confidence intervals indicate a true reduction between 0.3 and 1.5 pp. For GW the reduction is 1.8 pp, from 3.7 % to 1.9 % GW in the MW bin, or a 48 % reduction, with a true reduction estimated between 0.4 and 3.1 pp. The findings are in line with Syversen et al (2019), where the increase in collected metal is larger than the increase in collected glass when introducing a kerbside G&M collection. They found an increase in the amounts sorted at source of 11 % for glass and 157 % for metal. As metal packaging is lighter than glass packaging, we can assume that a smaller reduction in the share of metal in the MW can account for a larger increase in the amount collected from sorting at source.

Table 4 Results for simplified metal packaging waste fraction regression model.

Dependent variable: **Metal packaging waste**					
R^2	66 %	Adj. R^2	51 %	*95 % conf. intervals*	
Obs.	11	*Coeff.*	*P-value*		
(Intercept)		1.6 %	< 0.01	1.3 %	1.9 %
sepFW		0.7 %	0.07	-0.1 %	1.4 %
kerbG&M		-0.9 %	< 0.01	-1.5 %	-0.3 %
Undergr.		0.4 %	0.18	-0.2 %	0.9 %

For the MW, the results from the full regression model are shown in Table 6. The model explains (R^2) 90 % of the variation in correctly sorted MW in the MW bin. The only significant effect is

Table 6 Results for mixed waste fraction regression model.

Dependent variable: **Mixed waste**					
R^2	90 %	Adj. R^2	80 %	*95 % conf. intervals*	
Obs.	11	*Coeff.*	*P-value*		
(Intercept)		21 %	< 0.01	13 %	29 %
sepFW		18 %	0.01	6 %	29 %
kerbG&M		-2.7 %	0.66	-18 %	12 %
Undergr.		-0.02 %	1	-8 %	8 %
Covid19		9.1 %	0.16	-5 %	23 %
Year		-0.5 %	0.71	-4 %	3 %

Author Accepted Manuscript version of the paper by Tuva Grytli and Cansu Birgen
in Computer-aided chemical engineering, Vol 52 (2023) Page 2149-2154,
DOI: http://dx.doi.org/10.1016/B978-0-443-15274-0.50342-5

separate FW collection (p=0.01), which leads to an increase in the share of correctly sorted waste in the MW bin of 18 pp, from 21 % to 39 %, or an 83 % increase. The confidence intervals indicate a real increase in correctly sorted MW with separate FW collection of 6 - 29 pp.

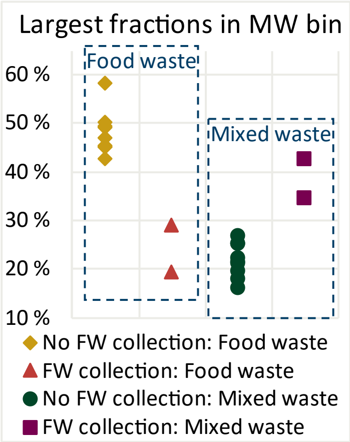

Figure 1 Shares of the two largest fractions in the mixed waste bin.

We see high explanatory powers (R^2) and large effects for both the FW and MW fractions, when separate FW collection is the significant variable. Two factors help explain this: (1) FW is heavy, constituting by far the largest fraction in the MW bin for households without separate FW collection, as shown in Figure 1. With separate FW collection, the correctly sorted MW is the largest fraction. (2) If there is no separate FW collection, the only option is to dispose of FW in the MW bin (or home composting). This implies that when we say *correctly sorted MW*, it does not mean that households without separate FW collection are doing anything wrong by discarding their FW in the MW bin, but rather that there should be separate FW collection, so this fraction can be recycled.

On the other hand, we saw how kerbside G&M collection did significantly affect the shares of GW and MeW in the MW bin, but not the share of correctly sorted MW. This is likely due to GW and MeW being small fractions in the MW bin. Effects may thus be obscured by variations in the larger fractions, such as food, plastic, and paper/cardboard. This shows the importance of looking at each fraction separately, to catch individual variations also in the smaller fractions.

Due to the weight of FW, we expected separate FW collection to impact the results of the shares of GW and MeW in the MW bin. As the heavy fraction is removed, the relative shares of all other fractions should increase. For MeW we saw this effect be close to significant, but we did not observe this for GW in the simplified model. The effect of the kerbside G&M collection was also less significant for GW in the simplified model. Through in-depth analysis of the background data, the reason for these results was found to be one outlier data point in a sample area after the rollout of kerbside G&M collection. The outlier lifts the sample area average above the lowest value before the rollout, and above the average of the areas with separate FW collection. This shows how the small number of sample areas can inflate the impact of noise from outliers. It is also a challenge when working with waste composition rather than absolute amounts. We found that by subtracting the FW from the waste composition, the problematic sample area's GW share falls below the lowest value before the rollout. This is because the earlier sample area had a very large share of FW, and the later (problematic) sample area a share below average, thus affecting the relative share of GW in the opposite direction. Both of these problems caused noise that made the simplified GW model seem overfitted, and as a consequence the choice to further simplify the model was made.

As discussed, the small number of observations caused some difficulties. However, using data from only one MSW company yields consistent data collection, which reduces other sources of uncertainty. Different methods for the waste composition analysis, in terms of for example sample size and location, or types and number of waste components, are good examples of such sources (Dahlén & Lagerkvist 2008). We assessed that the benefits from the data collection consistency outweigh the disadvantages of few observations and consider this study a good contribution to understanding the effects of external factors on the results of MSW sorting at source. To strengthen findings and make them more generalisable, similar analyses should be carried out for other MSW companies.

4. Conclusions

The regressions show that it is mainly separate sorting and user-friendly systems that influence the sorting results in the MW bin. The more fractions it is possible and easy to sort, the better the sorting results. We found no significant impacts over time, post covid-19 or from a different type of MW bin.

There are many factors that can influence sorting at source that we were not able to analyse in this study, such as demographic variables, communication, and information, which is further discussed in Grytli & Birgen (2023). To optimize the design of MSW systems for a circular future, it is crucial to identify and understand all relevant factors, and more research should be carried out to this effect.

This study has revealed challenges related to waste composition analyses due to manual data registration. As most MSW companies in Norway are small, the capacity for quality assurance is limited, increasing the risk of errors. Through this work we were able to provide debugging and data quality improvements for the MSW company. Improving the waste composition statistics will help the company better understand how their implemented measures are working and which new ones to consider.

Besides helping the specific MSW company, this study can help similar companies better understand and design their collection systems to increase circularity. We also aim to contribute to the necessary high-quality information and data needed for effective waste strategy planning by policymakers.

Acknowledgements

This work is part of the CircWtE project co-funded by industry and public partners and the Research Council of Norway under the SIRKULÆRØKONOMI program (CircWtE - 319795). The authors thank the personnel at the municipal solid waste company for providing data and valuable information on the waste management system.

References

L. Dahlén, A. Lagerkvist, 2008, Methods for household waste composition studies, Waste Management, Vol. 28, 7, 1100-1112, https://doi.org/10.1016/j.wasman.2007.08.014.

Directive (EU) 2018/851, 2018, Directive (EU) 2018/851 of the European Parliament and of the Council of 30 May 2018 amending Directive 2008/98/EC on waste (Text with EEA relevance)

T. Grytli, C. Birgen, 2023, Enabling circularity: Analysis of factors influencing MSW sorting behaviour in Central Norway, Proceedings of the 33rd European Symposium on Computer Aided Process Engineering (ESCAPE33), June 18-21, 2023, Athens, Greece. (Accepted)

Statistics Norway, 2022, Table 12263: Selected key figures for household waste (M) 2015 - 2021

F. Syversen, K. Kirkevaag, S. Bjørnerud, 2019, Renere råvarer. Case: Optimalisering av verdikjeden for glass- og metallemballasje, https://avfallnorge.no/fagomraader-og-faggrupper/rapporter/renere-r%C3%A5varer (accessed 2022-11-29)

J. Wooldridge, 2009, Introductory Econometrics: A Modern Approach, 4th Edition, Mason, Ohio: South Western Cengage Learning.