» Estimation of free riding in plastic package waste using put-on-market and business turnover information

Joao Lopes

Nuno Romao

Filipa Chambel

DMSI / ME

October 2023

Aim

» Calculate non-recycled plastic packaging waste

- Since 2020, EU member states contribute based on non-recycled;

- Non-recycled reported plastic packages put on market (POM);

- And undeclared plastic packages POM.

» Calculate undeclared plastic packages POM

- Annual business turnover (TO);

- Reported plastic packages POM.

» Establish relation between reported POM and TO

- $k = \dfrac{\text{POM}_{\text{reported}}}{\text{TO}_{\text{reported}}};$

- $\text{POM}_{\text{unreported}} = \text{TO}_{\text{unreported}} \times k.$

» Estimate total

- $\text{POM}_{\text{total}} = \text{POM}_{\text{reported}} + \text{POM}_{\text{unreported}};$

- $c = \dfrac{\text{POM}_{\text{total}}}{\text{POM}_{\text{reported}}}.$

» Procedure

1. Divide businesses in moderately homogeneous clusters;

2. Choose clusters with high-chance of free riding;

3. Define k for each chosen cluster.

» Main assumption

- Ratio k capture characteristics of groups of businesses.

Data > data source

» From Statistics Portugal

- Business TO from sales of goods and sales of products;

- Number of employees (NOE);

- Economic activity classification with 5 levels (CAE3).

» Limitations

- Data can be updated/corrected by businesses after retrieving;

- Data goes through checks and validation after retrieving;

- Only businesses located in Portugal (and not full universe).

Data > data source

» From APA

- Municipality POM via Producer Responsibility Organizations (PRO);

- Non-municipality POM directly via producers.

» Limitations

- New clients adhere retroactively;

- Data can be updated/corrected by businesses after retrieving;

- Smaller businesses do not report values (use of estimates);

- Data from producers not consolidated.
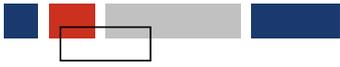
Data > data merging

» **Create final dataset**

- Secondary CAE3 reported as main activity;

- Sum over sale of goods and sale of products as total TO;

- Sum over municipality and non-municipality POM.

ID	NAME	YEAR	CAE3	NOE	POM	TO	k
999999990	business1	2014	47111	201	152	3512	0.04
999999990	business1	2015	47111	203	164	3812	0.04
...	...	...	...	...	...	...	...

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

Free riding in plastic package waste
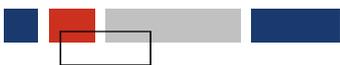
Methods > business clusters

» **Cluster characterization**

- Main variables: POM, TO, k;

- Summary statistics: $n, t, \hat{m}, sd, CV$.

» **Creating clusters**

- Economic activity classification (CAE3): A, C, G, I, N;

- Number of employees (NOE): micro, small, medium, large.
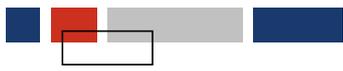
INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL

Free riding in plastic package waste

Methods > business clusters

» Criteria

1. Representativeness in numbers of business $[n_\text{TO}'/(n_\text{TO} + n_\text{TO}')]$;

2. Representativeness in business turnover $[t_\text{TO}'/(t_\text{TO} + t_\text{TO}')]$;

3. Robustness of the estimate of k $[n_k]$;

4. Homogeneity of clusters regarding k $[CV_k]$.

» Selecting clusters for estimation (via thresholds)

- Aggregating/Disaggregating clusters;

- Discarding clusters with low-chance of free-riding.

Methods > business clusters

» Criteria

1. Representativeness in numbers of business $[n_\text{TO}'/(n_\text{TO} + n_\text{TO}')]$;

2. Representativeness in business turnover $[t_\text{TO}'/(t_\text{TO} + t_\text{TO}')]$;

3. Robustness of the estimate of k $[n_k]$;

4. Homogeneity of clusters regarding k $[CV_k]$.

» Naive first approach

- Educated *ad hoc* method (36 clusters);

- Threshold for free riding ($c = \text{POM}_\text{total}/\text{POM}_\text{reported} < 5$).

» Different approaches

- Ratio of summary statistics $k_{totals} = \frac{\sum_i^n \text{POM}_i}{\sum_i^{n'} \text{TO}_i}$;

- Summary statistics of ratio $k_{mean} = \sum_i^n \frac{\text{POM}_i}{\text{TO}_i}/n = \sum_i^n k_i/n.$

» Advantage of summary statistics of ratio

- Analyse ratios at microdata level;

- Inform about quality of estimates (e.g. confidence intervals).

» Disadvantage of summary statistics of ratio

- Sensitive to large discrepancies.

- Complete data on municipality and non-municipality POM;

- No overlap between municipality and non-municipality POM;

- Value of goods only from plastic packages;

- **Non-declarants are free riders**;

- **Homogenization within clusters of businesses**;

- Complete data on TO of national businesses;   overestimation

- Negligible secondary economic activity;

- No free riding in discarded clusters;

- **Complete data on TO of foreigner businesses**.   underestimation
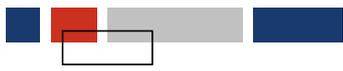
Exploratory results > tables

» Description of POM (10^3 t), TO (10^6 €) and k (kg/€)

CAE	n_{report}	POM$_{report}$	TO$_{report}$	k_{means}	$n_{unreport}$	TO$_{unreport}$
A	852	1.0	829	0.017	126796	6,976
C	3779	854.3	61,184	0.014	66662	34,694
G	3625	132.6	64,325	0.004	221877	87,832
I	148	0.6	192	0.006	115858	1,039
N	68	4.3	116	0.096	187655	1,232
other	384	1.2	5,098	-	664463	33,613
unavailable	358	2.0	-	-	-	-
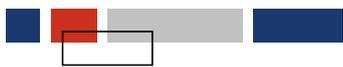
Exploratory results > plots

» Estimates of POM and TO

» Estimates of *k* and proportion of declarants
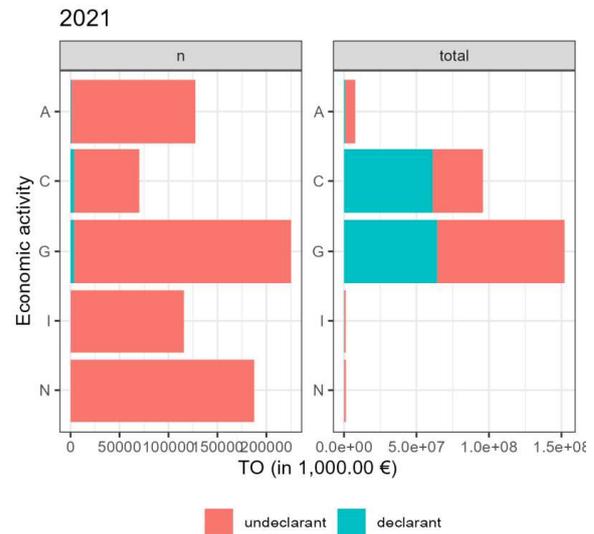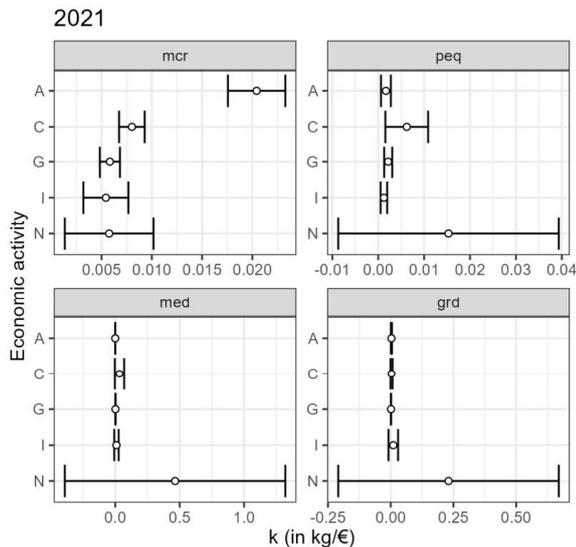
Future work

- Preserving data on non-municipality POM;
- Using robust summary statistics;
- Disaggregating clusters of businesses;
- **Choosing clusters of businesses;**
- **Choosing values of thresholds (sensitivity analysis);**
- Filtering out businesses for calculating ratios k;
- Accounting for discarded clusters of businesses;
- Other future developments.

Thank You

Annexes
⌄

» Categories of economic activity

	description	2 digits	3 digits	4 digits	5 digits
A	Agriculture, forestry and fishing	3	13	39	49
C	Manufacturing	24	95	230	332
G	Wholesale and retail trade…	3	21	91	124
I	Accommodation and food…	2	7	9	34
N	Administrative and support…	6	19	33	35

» Categories of number of employees

	description	definition
mcr	Micro-size business	NOE $\leq$ 10 employees
peq	Small-size business	11 employees $<$ NOE $\leq$ 50 employees
med	Medium-size business	51 employees $<$ NOE $\leq$ 250 employees
grd	Large-size business	NOE $>$ 250 employees

Data > data merging

» Main problem

- Prior 2017: reports on non-municipality POM by PROs;

- 2017 onwards: reports on non-municipality POM by producers;

- Data consolidation not performed in reports by producers.

» Solution

- Outlier diagnosis and imputation (before merge) – check integrity;

- Outlier diagnosis and imputation (after merge) – soften transition.
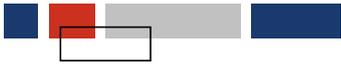
Data > data merging

» Internal outlier diagnosis (across time)

- $x > Q75 + 3 \times IQR;$

- $x > 10 \times Q50.$

» External outlier diagnosis (across businesses)

- $x > Q75 + 3 \times IQR;$

- $x > Q99.$
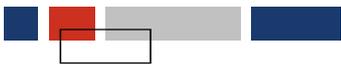
Data > data merging

» Categories of outliers

- <u>scalar</u>: values reported in Kg (instead of tonnes) - scale down value;

- <u>replace</u>: clerical errors - replace value by median per business;

- <u>ignore</u>: incorrect labelling as non-municipal - replace value by 0;

- <u>keep</u>: incorrect identification as outlier - keep value.
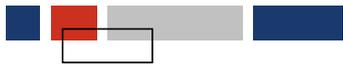
» Description of non-household POM (10^3 ton)

Outlier treatment	2021	2020	2019	2018	2017
before	42.5	46.3	48.9	55.3	49.0
after	37.7	37.4	39.1	37.8	33.1

» Description of non-household POM (% total POM)

Outlier treatment	2021	2020	2019	2018	2017
before	22%	25%	27%	31%	29%
after	20%	20%	22%	21%	20%

INSTITUTO NACIONAL DE ESTATÍSTICA
STATISTICS PORTUGAL