Accompanying document for REFRESH D1.4, "Quantified consumer insights on food waste".

This document provides background information on the recruitment and sampling strategy that was used.

Sampling problem

The design was meant to have an identical number of sampling units (N=1000) in each of the countries, using a quota sampling approach per country. Data were collected through a renowned international market research agency (GfK) through its Dutch branch (GfK Amstelveen). GfK subcontracts the data collection to its national partners, in this case in Spain, Germany, the Netherlands and Hungary.

Unfortunately, due to miscommunication, an inconsistency in the setting of the desired sampling quota appeared across the countries. In two countries (Germany and Hungary) the national quota were erroneously set at the individual level, whereas in two other countries (Spain and the Netherlands) these were correctly set at the household level. In other words, in the Netherlands and Spain the quotas were calculated by dividing the national percentage of each household size (e.g., 1 person household) relative to the total amount of households in that country. In Germany and Hungary quota for each type of household size (e.g., 1 person household) were mistakenly expressed relative to the total number of inhabitants of that country.

The lack of representativeness of the sample has been reported at the plenary REFRESH General Consortium Meeting in Budapest (7-8 September 2017). After identifying the problem as one of inconsistently defined sampling quota, there has been a constructive dialogue between the REFRESH team and the market research agency, as well as with the Wageningen University Social Science Ethical committee and the Wageningen University Research Integrity Committee, to explore the most appropriate solution.

Chosen solution

After careful consideration of the pros and cons of the different options, we have selected one that balances the consideration of (a) effective use of data, (b) representativeness (and comparability across countries) of the data, (c) value for money to the commissioner, and (d) economic/financial damage to the market research agency. This solution involves the following:

The market research agency randomly sampled cases from the German and Hungarian samples, such that the resulting subsample reflected the household level quota. This was done randomly with known starting values to allow for replicability and ensure complete transparency of the procedure. All data collected in Spain and the Netherlands, which was collected according to intended quota, was retained.

Alternative solutions were rejected for various reasons. Weighing the data from Germany and Hungary to make the samples fit household level quota to which these samples were not designed would involve problematically high weighing factors. Collecting additional data according to the correct quota would lead to issues due to the seasonal dependency of the topic of research (household food

waste is known to vary by season). Ignoring the data for the two countries would seriously reduce the value for money for the commissioner.

The outcome

As a result of this approach, the samples used in this research are the following:

- Original collected sample for Spain (N = 1052) and the Netherlands (N = 1054), even though these samples do no fully concord to the set quota for these countries, as there is always a deviation from the quota due to the nature of data collection.
- In Germany, the realisation diverged considerably from the individual level quota. As a result, a fairly efficient sample at household level occurred by random selection of cases to make it fit with household level quota, resulting in a final sample of N = 872.
- In Hungary, the realisation was more close to the a-priori set individual level quota, and to match it to the household level required a considerable loss of cases, resulting in a final household level sample of N = 471.

The table below shows the original and final sample characteristics for the four countries in terms of household size.

Household size distribution	Eurostat population distribution	Raw sample	After filtering
Germany		N = 1053	N = 872
1 Person household	40.7%	33.7%	40.7%
2 Person household	34.2%	38.0%	34.2%
3 Person household	12.4%	15.9%	12.4%
4 Person household	9.3%	9.0%	9.5%
5+ Person household	3.4%	3.4%	3.2%
Hungary		N = 966	N=471
1 Person household	33.7%	16.4%	33.7%
2 Person household	30.4%	34.4%	30.4%
3 Person household	16.7%	23.4%	16.8%
4 Person household	12.4%	16.3%	12.4%
5+ Person household	6.9%	9.6%	6.9%
Spain		N = 1052	n/a
1 Person household	25.4%	10.4%	
2 Person household	30.4%	22.9%	
3 Person household	20.9%	28.2%	
4 Person household	17.4%	30.8%	
5+ Person household	5.8%	7.7%	
The Netherlands		N = 1054	n/a
1 Person household	37.7%	43.7%	
2 Person household	32.8%	32.0%	
3 Person household	12.0%	9.6%	
4 Person household	12.4%	10.3%	
5+ Person household	5.2%	4.4%	

Note: Eurostat data is used as a reference (data obtained at 10.2017).

http://ec.europa.eu/eurostat/web/income-and-living-conditions/data/database