Consumers behavioural economic interrelationships and typologies

REFRESH is funded by the Horizon 2020 Framework Programme of the European Union under Grant Agreement no. 641933. The contents of this document are the sole responsibility of REFRESH and can in no way be taken to reflect the views of the European Union

Authors

Matthew Grainger, University of Newcastle

Gavin Stewart, University of Newcastle

With contributions by:

Lusine Aramyan, Wageningen Economic Research, The Netherlands

Natalia I. Valeeva, Wageningen Economic Research, The Netherlands

Tom Quested, WRAP, UK

Erica van Herpen, Wageningen University, the Netherlands

Simone Piras, University of Bologna, Italy

Simone Righi, University of Bologna, Italy

Marco Setti, University of Bologna, Italy

Matteo Vittuari, University of Bologna, Italy

Project coordination and editing provided by Wageningen University.

Manuscript completed in June, 2016

This document is available on the Internet at: http://eu-refresh.org/

Document title	Identification of different consumers behavioural, economic inter-relationships and typologies
Work Package	WP4
Document Type	Deliverable
Date	June 30th 2016
Document Status	Final version

Acknowledgments & Disclaimer

This project has received funding from the *European Union's Horizon 2020 research and innovation programme* under grant agreement No 641933.

ISBN 978-94-6257-989-7

Executive summary

The reduction of food waste has become one of 17 established United Nations Sustainable Development Goals. There is now an international target of halving per capita food waste at the retail and consumer level and reducing food losses along production and supply chains by 2030. Consumer food waste in Europe has been estimated at between 6 % and 15 % of food purchased, representing greater loss than retail or production supply chain losses. Understanding consumer behaviour in relation to food waste generation is therefore critical in developing and targeting cost-effective interventions to achieve food waste reductions.

Quantifying and understanding inconsistency in the generation of consumer food waste is particularly important as identification of consumer waste typologies allows targeting of intervention strategies where they are most efficient. The current literature identifies large numbers of characteristics which identify typologies, but many of these relationships are likely to be spurious, as a result of multiple testing, small sample sizes, selective reporting and lack of pre-specification of hypotheses limiting the proportion of true to untrue relationships under investigation.

We addressed these problems by analysing the two largest available datasets using a range of analytical techniques that minimise type I (false positive) errors. Use of data from WRAP had the further advantage of allowing comparison of self-reported waste with measured waste; whilst Euro-barometer data, facilitated generalisation across Europe.

We analysed both data sets using multiple regression models. Rather than specifying a single model, we accepted that different model structures could lead to different results and used model averaging to account for this uncertainty therefore reducing the false-positive error rate. We also utilised machine learning methodologies (random forests and Bayesian Networks) to corroborate the findings.

Analyses consistently indicated that consumers (households) were variable in their food waste behaviour reinforcing the importance of identifying typologies. They also consistently identified ***household composition as a key typology or determinant of food waste*** with large households generating more waste than small households. Households should therefore be considered as an important unit of analysis in further work, although this does not preclude further exploration of individual actors as well. Analyses of Euro-barometer consistently identified ***Country*** as an important predictor with some indication that countries with grocery spending per capita in excess of €3000 had higher food waste but no apparent relationship with GDP (these relationships will be explored further in future work to ascertain how robust they are). Demographics, specifically ***age, education*** and ***occupation*** were important predictors in some analyses but not others, and are therefore tentative candidate variables for consumer waste typologies.

In addition to identifying typologies, we were also concerned with generating systems map to allow holistic modelling of waste as an emergent property of a complex system. Consumer behaviour is complex with large numbers of drivers, effect modifiers or context dependencies, and behaviours interacting. Identifying commonalities and developing systems models of consumer behaviour is therefore a pre-requisite for developing accurate predictive models to inform policy.

We used machine learnt Bayesian Networks to develop systems maps of the consumer food waste nexus. Different linkages were emphasised to different degrees in models based on different data but one important commonality emerged. ***Consumer behaviour before shopping, in the retail environment,*** and ***in the home predicted food waste.*** Modelling of consumer behaviour should therefore not be restricted to a single environment. The clustering of behaviours and drivers within each environment requires further unpacking to ascertain which behaviours are key in which environments.

The strength of evidence underpinning the generation of typologies and systems map is currently low. Analysis of the WRAP data indicates that a large proportion of people who self-report low or no waste, actually waste. This indicates that outcomes relying solely on self-reporting are at high risk of bias potentially underestimating waste generation and confounding relationships between behaviour and waste generation where biases differ systematically. The Eurobarometer data indicates large country differences in waste generation which remain poorly understood. Understanding this heterogeneity is key in further generation of typologies. Precision is generally low where specific relationships are examined and this is exacerbated by problems with multiple testing. We did not explicitly evaluate the evidence base for publication or selective reporting biases as analyses were based on raw data, but we highlight the pervasive nature of these biases. Further data collection can ameliorate these problems, particularly in relation to precision and inconsistency; but the uncertainty inherent in information on consumer behaviour requires appropriate propagation in probabilistic models, to inform coherent decision-making.

Table of Contents

List of Tables

List of Figures

List of Boxes

List of abbreviations

BN Bayesian Network

WRAP The waste and resources action partnership a UK registered Charity (No. 1159512) see www.wrap.org.uk

d.f. Degrees of freedom

AICc Akaike Information Criterion corrected for small sample size

EU European Union

1 Background, objectives, relevance

1.1 Background information

Consumer food waste in Europe has been estimated between 6 % and 15 % of food purchased (Euro-barometer 2013). This estimate comes from self-reported surveys across 28 EU countries. Waste produced at the post-consumption stage is thought to be responsible for the largest proportion of all food waste in developed countries (Parfitt et al. 2010). Stenmarck et al. (2016) estimated food waste in the 28 EU countries (extrapolated from data for 11 countries) at 88 ± 14 (95% CI) million tonnes with 47 ± 4 (95% CI) million tonnes coming from households. This equates to between 47% and 64% of total European food waste coming from households.

Identified drivers of consumer food waste

Consumer food waste has been defined and measured in a variety of ways leading to difficulties in synthesis (Sibrián et al. 2016; Møller et al. 2013). For example, it has been measured through direct measurement (e.g. Wenlock & Buss 1977; Quested & Luzecka 2014; Parizeau et al. 2015) and estimated through food diaries or other self-reporting mechanisms (e.g. van Garde & Woodburn 1987; Koivupuro et al. 2012; Stefan et al. 2013; Abeliotis et al. 2014). Factors that have been identified as important in determining the levels of household food waste include (Bos-Brouwers et al. 2012; Canali et al. 2013):

- Demographics – e.g. household size (Wenlock & Buss 1977; Barr 2007; Koivupuro et al 2012; Quested & Luzecka 2014; Parizeau et al. 2015), household composition (i.e. age structure and gender of main shopper) (Wassermann & Schneider 2005; IGD 2007; Glanz 2008; Koivupuro et al. 2012; Quested & Luzecka 2014; Parizeau et al. 2015), employment status (Wassermann & Schneider 2005), income (Wenlock et al. 1980; Muth et al. 2007; Baker et al. 2009; Parfitt et al. 2010; Katajajuuri et al. 2012) and education level achieved (Wasserman & Schneider 2005; Silvennoinen et al. 2012).
- Consumer knowledge – e.g. understanding of date labels on products (van Garde & Woodburn, 1987; IGD 2007; WRAP 2007; WRAP 2011; WRAP 2012; Quested et al. 2013; Abeliotis et al. 2014).
- Consumer preferences – e.g. Preference for perceived high quality food items (WRAP 2011; Sonesson et al. 2005).
- Consumer behaviour – e.g. use of leftovers (WRAP 2007; WRAP 2008, Stancu et al. 2016); Frequency of shopping (Koivupuro et al. 2012; Quested et al. 2013; Stefan et al. 2013; Stancu et al. 2016).
- Consumer attitudes – e.g. knowledge of waste and recycling (Barr 2007; WRAP 2007; Quested & Luzecka 2014; Parizeau et al. 2015; Stancu et al. 2016).
- Social norms – e.g. not storing fruit in the fridge so that it can be on display (Johnson et al. 2009; George et al. 2010).

- Consumer tools – e.g. access to transport and/or local shopping opportunities (WRAP & French-Brooks 2009; Silvennoinen et al. 2012).

What is clear from the literature is that food waste occurring in the home (consumer waste) has multiple potential drivers that maybe inter-related. Demographic factors such as household size (number of people in the house) and household composition (the age/relationship structure in the house) are identified in many studies as being important in determining food waste volumes. These are most likely proximate factors that are determining behaviours that lead to food wastage in the home.

Types of food waste

The main types of food waste recorded in the literature constitute lower cost products such as bread, milk, rice and pasta regardless of the study country (e.g. Pekcan et al. 2005; IGD 2007; Glanz 2008; Sonesson et al 2005; Quested et al. 2011). However, measurements of "importance" differ across studies, some addressing the weight, some estimating a percentage of the total waste and others estimating the calorific value of food types disposed.

The amount of fruit and vegetable wastage (and to a certain extent meat waste) appears to be driven in part by the discard of inedible parts (stones, skins, cores, etc.; Wassermann & Schneider 2005; IGD 2007; WRAP 2008). In the UK nearly half of estimated 1.9 million tonnes of fresh vegetable and salad waste is categorised as "unavoidable/possibly unavoidable waste" (WRAP 2009). WRAP (2009) define these food waste categories as follows:

- "Avoidable – food and drink thrown away that was, at some point prior to disposal, edible (e.g. slice of bread, apples, meat).
- Possibly avoidable – food and drink that some people eat and others do not (e.g. bread crusts), or that can be eaten when a food is prepared in one way but not in another (e.g. potato skins).
- Unavoidable – waste arising from food or drink preparation that is not, and has not been, edible under normal circumstances (e.g. meat bones, egg shells, pineapple skin, tea bags)."

Most analyses of consumer food waste data are limited because we do not know the weight of purchased food. Large households waste a greater amount of food than smaller households but we might expect that the large households will also purchase a larger amount of food (although household size is still important even when controlled for and thus may be interacting with other variables; Quested & Luzecka 2014). Self reporting waste as a proportion of that purchased is one approach used to overcome this (e.g. Secondi et al. 2015; Stancu et al. 2016). Adelson et al. (1963) is, to our knowledge, the only study to have weighed both purchased and consumed food. They weighed food purchased in volunteer households in three locations in the USA (see Table 1) and then weighed waste in 4 food categories (meat/poultry/fish; dairy products; grains/sugar; vegetables).

Table 1. Recreation of Table 1 from Adelson et al. (1963) showing the difference between food purchased and food wasted in the household (calculated by calorific value).

			Site 1	Site 2	Site 3
		Year of Study	1959	1958-59	1959-62
		Number of Households	60	62	64
		Range of household size	3.5 to 3.7	4.5 to 4.6	3.4 to 3.5
Food energy	As purchased	Calories	3025	2550	2570
	Lost amount	Calories	245	175	250
	Lost proportion	Percent	8	7	10
	Presumed eaten	Calories	2780	2375	2325
Nutrient fat	As purchased	Calories	1300	1070	1110
	Lost amount	Calories	135	115	200
	Lost proportion	Percent	11	11	18
	Presumed eaten	Calories	1160	955	915
Calories from fat - Total calories					
	As purchased	Percent	43	42	42
	Presumed eaten	Percent	42	40	38

Although there are increasing numbers of studies on the drivers and types of food waste it is difficult to develop consumer typologies solely from the literature because of high inconsistency and variation in methodology. Further analysis was undertaken to develop behavioural typologies and systems maps to inform future modelling of consumer behaviours in relation to food waste

1.2 Objectives

Sub-task 4.1.2 is a part of Task 4.1 ("Socio-economic implications of food waste"), which aims at identifying and measuring the major socio-economic conditions and driving factors that influence business and consumer choice in the creation or reduction of food waste. The objectives of sub-task 4.1.2 are to:

1. Define consumer behavioural typologies on the basis of relevant literature and assessment of datasets relating waste to consumer behaviour?

2. Develop a systems map to illustrate potential *links between consumer behaviour and the creation / reduction of food waste.*

2 Research methodology

There are limited data available on consumer behaviour in relation to food waste (see WP1 Literature Review). Studies on food waste use a variety of approaches but most can be divided in two broad categories; those relying on some form of self-reporting or those relying on some form of objective measurement (i.e. weight and composition of waste). Here we use data the two largest (most precise) studies identified from the WP1 Literature Review. One is a geographically limited dataset (limited to England and Wales) which uses an objective measurement of food waste and the second is a dataset with a wide geographic extent (the 28 countries of the EU) which uses self-reported estimates of food waste. Additionally we explore data on European Municipal waste collections for a subset of the 28 countries of the EU. The nature of these datasets means that we can investigate two sub-objectives related to Objective 1. These are to:

1a) Identify if an objective outcome (i.e. weight and composition of waste) differs to self-reporting waste and,

1b) Assess the generalisation of food waste data from a single country to the whole of the EU.

2.1 Datasets

Dataset 1: WRAP's "Waste compositional analysis" and "Kitchen Diary 2012" (provided by T Quested at WRAP).

Mostly the "Waste compositional analysis" dataset is utilised herein. This consists of face to face in home interview responses (categorical data) on social/demographic aspects of households, behavioural responses to food waste along with data on the amount of waste collected from the kerbside. The sample size was 1799 UK households. Waste was collected by teams from outside each home (flats and houses with shared waste collections were not assessed). After collection, waste from each household was weighed and sorted. All non-food items were removed and weighed. Food items without packaging were sorted by food type and then weighed. For food items with packaging, these were removed from the packaging, weighed separately and any details on the packaging (e.g. best before dates) were recorded. For more details see Quested et al. (2013) and references within.

Dataset 2: Flash Euro-barometer 388: Attitudes of Europeans towards waste management and resource efficiency (available at http://open-data.europa.eu/en/data/dataset/S1102_388)

A sample of 26595 European citizens (aged 15 and over) were asked questions on the telephone about their opinions and behaviours in relation to food waste management and resource efficiency in December 2013. Samples were taken in 28 EU countries with a mean sample size of 949.71 ±158.41 (mean ± standard deviation). However, three smaller (in terms of population size; Republic of Cyprus, Luxembourg and Malta) countries had a mean sample size of 501 ±1. The remaining 25 countries had a mean sample size of 1003.60 ± 6.59. A total of 20 questions related to food waste and resource efficiency and a further 22 questions in relation to demographics (household size, employment, education, sex, etc.) were asked of each respondent. For technical details and a list of questions asked during the survey please see http://ec.europa.eu/public_opinion/flash/fl_388_en.pdf.

Dataset 3: OECDs municipal waste dataset (see Bagherzadeh et al. 2014 and http://www.oecd.org/env/waste/oecdenvironmentaldatacompendium.htm). Municipal waste includes solid waste from households as well as some business and commercial properties that is collected by the local municipality. Data were collected from government and academic sources (websites), with an additional request for data sent to members of the OECD Working Party on Environmental Information of the Environment Policy Committee. This included questions on the tonnage of food waste generated. Data on municipal waste per country were only available for 11 European countries (however we were able to extrapolate to other countries see below for details).

2.2 Data analysis

Robust analysis is dependent on multiple methods providing confirmation of results. Three types of analytical method that allow the identification of the "importance" of variables in explaining food waste generation (both self-reported and objectively measured) were used. These methods were regression analysis (Generalised Linear Models, GLM) and two machine-learning algorithms (Random Forest and Hill-climbing) to develop regression and classification trees as well as Bayesian networks. Full analytical details are available in the technical appendix. In summary, GLM is a flexible generalisation of ordinary linear regression where the relationship between a response variable and a number of predictors can be assessed. Model structure (i.e. the number and identity of predictor variables used) can result in misleading results and as such model-selection is used to select the most parsimonious model (i.e. a model that accomplishes a desired level of prediction with as few predictor variables as possible) in all possible combinations of predictors. Due to the large number of possible combinations of predictors the Random Forest algorithm was used for model reduction in the regression analysis. In addition, we removed highly correlated predictor variables as these can falsely inflate the coefficient and standard errors (i.e. the size of the effect).

Machine learning is a subfield of computer science that is related to the study of pattern recognition and artificial intelligence. The two algorithms used are designed to recognise relationships between variables and to show how important each variable is to the response (in this case food waste). Random Forests creates many regression and classification trees (minimum 500) where the data are split in to different branches of the tree to best explain the response variable. Bayesian networks are a graphical representation of a network of variables whereby related variables are joined by an arc (or arrow) and a set of conditional probabilities (where the state of one variable is conditional on the state of another). Machine-learnt Bayesian networks can recognise relationships between variables but not the direction of the relationship so arrow heads are added at random. Machine learning is much more robust to highly correlated variables than regression analysis.

A schematic of the analysis process linked to each objective and sub-objective is shown in Figure 1.

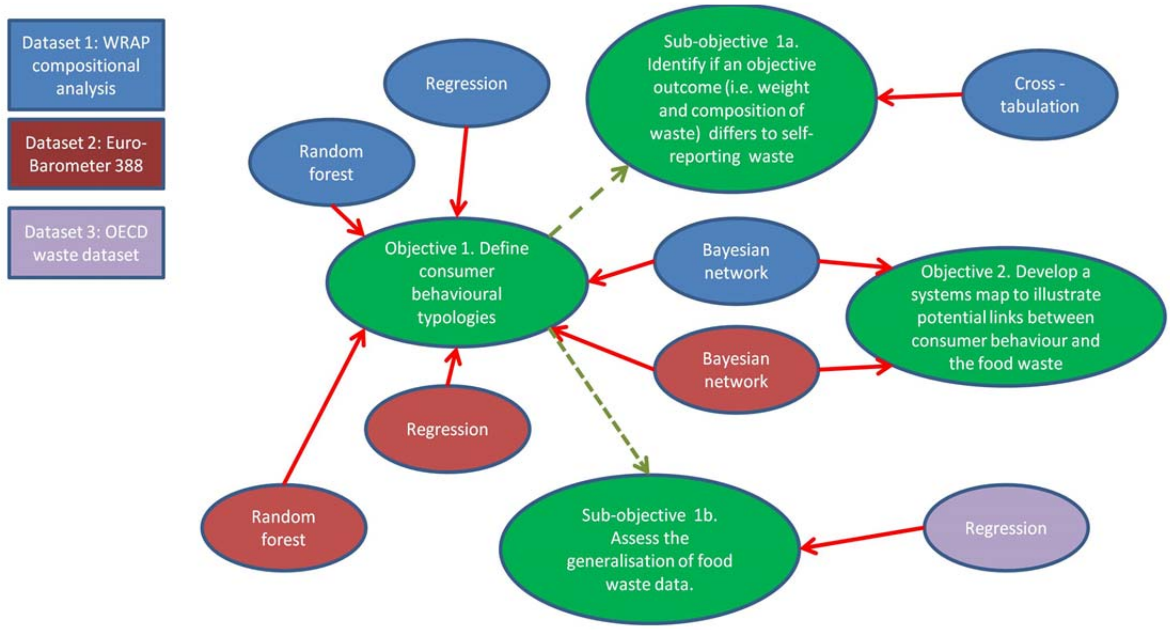


Figure 1. A schematic representation of the methodological approach. For full methodological details please see the technical appendix.

3 Results

Full results are available in the technical appendix. What follows is a summary of the results ordered by the objectives.

3.1 Objective 1: Define consumer behavioural typologies on the basis of relevant literature and assessment of datasets relating waste to consumer behaviour

For dataset 1 (WRAP dataset) household size was consistently identified as being an important driver of avoidable food waste (even when "controlled for" in the regression models) in all three analyses and was highly ranked in each (Table 2). Discarding behaviours were also highly ranked across all analyses, but the food type discarded differed. The random forest identified shopping and planning behaviours as being important, although they were not identified as important in the other analyses.

Table 2. A comparison of important drivers of avoidable food waste for dataset 1 (WRAP dataset) across the three analytical approaches ranked within each analysis[1].

Regression	Random Forest	Bayesian Network
Discard vegetables	Use of leftovers	Household size
Household size	Pre-planning & list writing	Household composition
Local authority	Discard parts	How waste is disposed
Discard bread	Household size	Presence of fussy eaters
Discard parts	Discarding meals	Discarding meals
	Looking in cupboards for vegetables	Discard cooked
	Discarding vegetables	Discard packs
	House type	Discard sell by date

[1]Variables shared between all analyses are shaded green, those shared between two analyses are shaded red and those only identified in one analysis are shaded blue

For the Euro-Barometer 388 dataset country, age of respondent, education and a self-reported belief that the family wasted too much were consistently important in all three analyses (Table 3). Household size and household preference for waste management were important variables in the regression and random forest. The random forest identified agreement that employment opportunities stem from waste management, local litter levels and community type as being important, although they were not identified as important in the other analyses. Occupation was identified as important in the Bayesian network but not in the other analyses.

Table 3. A comparison of important drivers of self-reported food waste for dataset 2 (Euro-barometer 388) across the three analytical approaches ranked within each analysis[2].

Regression	Random Forest	Bayesian Network
Age	Country	Age
House management	Age	Country
Country	Household size	Education
Education	House management	Occupation
Too much waste	Employment opportunities	Too much waste
Household size	Litter	
	Too much waste	
	Economic growth	
	Community type	
	Education	

[2]Variables shared between all analyses are shaded green, those shared between two analyses are shaded red and those only identified in one analysis are shaded blue

Bayesian networks allow users to set scenarios and return the probability that variables will be in a certain state. For dataset 1 (WRAP) and dataset 2 (Euro-barometer 388) the response variables (avoidable food waste and self-reported food waste respectively) were set at the highest and lowest state reflecting the maximum and minimum quartile of waste (Figure 2 & Figure 3) to help refine typologies. For the WRAP dataset (1) it was difficult to determine clear typologies (Figure 2). Further analysis excluding discard behaviours may lead to additional understanding but the magnitude of the signal to noise ratio inherent in the data is likely responsible. In general at the lowest levels of food waste the distribution of probabilities are more skewed to the right than at the highest levels of waste. This can be interpreted as smaller household sizes are more likely to produce lower levels of waste, but there is a large amount of variation in waste generation that cannot be explained by household size (Figure 2). Food waste was likely to be lower when it was collected by the council in the general waste bin (Figure 2).

Typologies were clearer to determine for the Euro-barometer dataset (2) (Figure 3); the age node was more likely to be in an older state when the scenario was set at the lowest level of waste. This can be interpreted as older people self-report less waste than younger people. The education node was more likely to be in the ‘still-studying’ state when the scenario was set to the highest level of waste (i.e. students are more likely to create higher levels of waste – greater than 50% of purchased food; Figure 3). Reporting that a household creates too much waste corresponded with the high waste scenario as would be expected (Figure 3).

3.2 Objective 1a: Identify if an objective outcome (i.e. weight and composition of waste) differs to self-reporting waste

Dataset 1 (WRAP) used an objective measure of waste (the weight of avoidable food waste) alongside some self-reported measures of waste type and frequency of discard. The question “Over the last week, how much of the following foods have you thrown away…” was asked for a variety of foodstuffs with possible answers including “Quite a lot”, “A reasonable amount”, “Some”, “A small amount”, “Hardly any”, “None”, “Don't eat it” and “Not stated”. A cross tabulation was used to assess the frequency of these answers for high, medium or low levels of avoidable food waste.

There were some apparent inconsistencies between self-reported discard behaviour and the amount of avoidable food waste produced by survey respondents (Figure 4). Few respondents reported discarding “quite a lot” or “a reasonable amount” of food waste (with the exception of inedible parts of food and waste left on plates after meals). There is some evidence that reporting that ready meals and bread are not eaten leads to less probability of being in the highest waste category (which agrees with the regression analysis).

This finding suggests that self-reported levels of waste are not reliable when respondents are asked to estimate food waste.

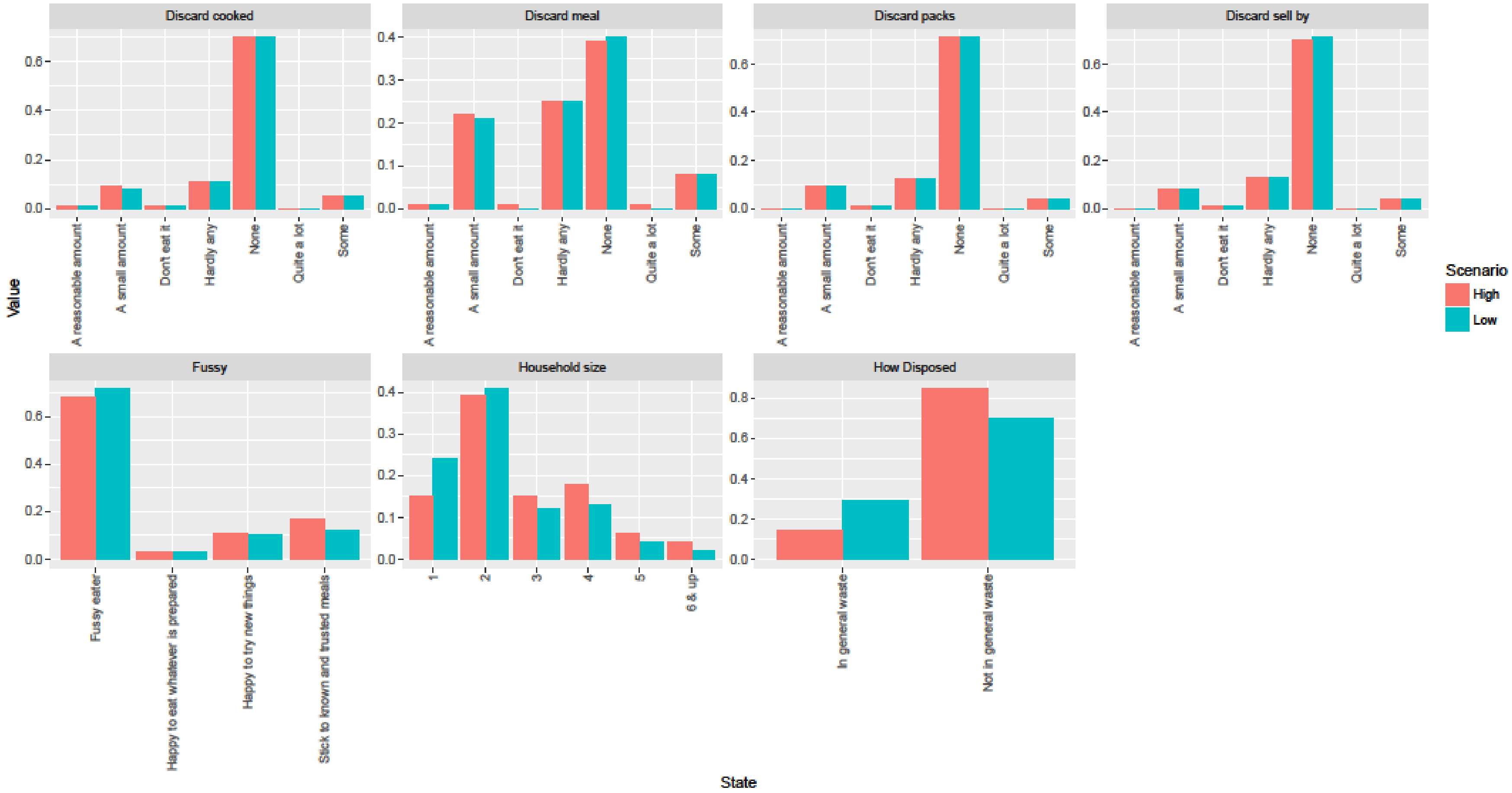


Figure 2. The probability of each node state for the discard cooked, discard meal, discard packs, discard sell by, presence of fussy eaters, household size and how waste is disposed nodes of the WRAP (dataset 1) Bayesian network. Avoidable Food waste is set at the highest and lowest quartile in the scenarios compared.

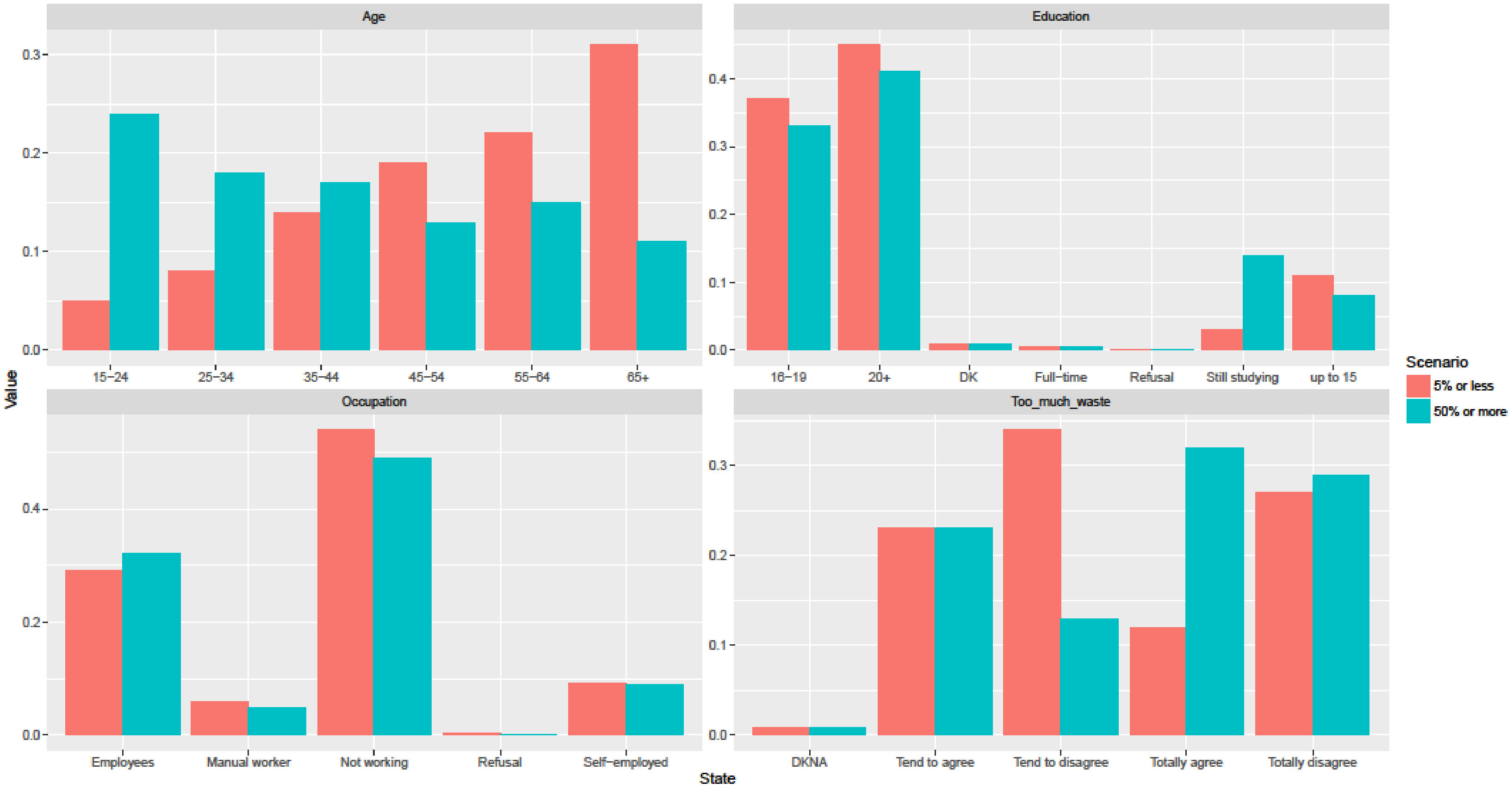


Figure 3. The probability of each node state for the age, education, occupation and too much waste nodes of the Eurobarometer 388 (dataset 2) Bayesian network. Food waste is set at 5 % or less and at 50 % or more of purchased food.

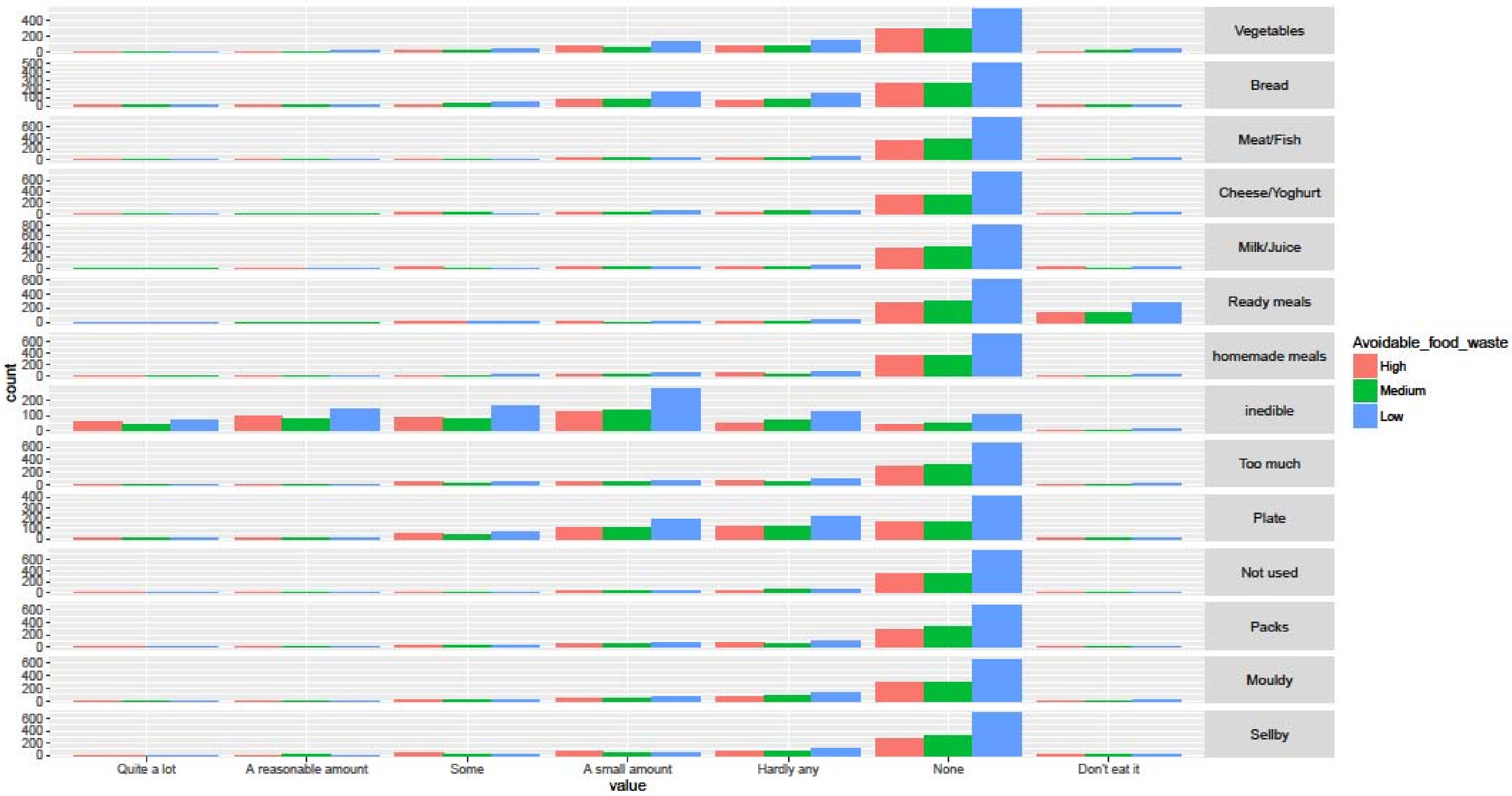


Figure 4. Outcome of question on discard behaviour divided by the level of food waste

3.3 Objective 1b: Assess the generalisation of food waste data from a single country to the whole of the EU

The Euro-barometer 388 dataset (dataset 2) allowed for the investigation of generality from a single country to the whole of Europe. The analysis of the Flash Euro-barometer 388 above (for Objective 1) highlighted country-level differences in self-reported food waste. The majority of respondents in the majority of countries report 5 % or less of purchased food going to waste. Where differences between countries are apparent is in the amount of respondents reporting higher levels of waste (Figure 5).

Figure 5. Self-reported percentage of purchased food wasted by European consumers (from the Euro-barometer 388 dataset).

To investigate potential causes of these differences the OECD waste dataset (dataset 3) was utilised. Data on municipal waste per country were only available for

11 countries. However, municipal waste was positively correlated with population size, allowing prediction (using regression analysis) for the remaining 17 countries (Figure 6).

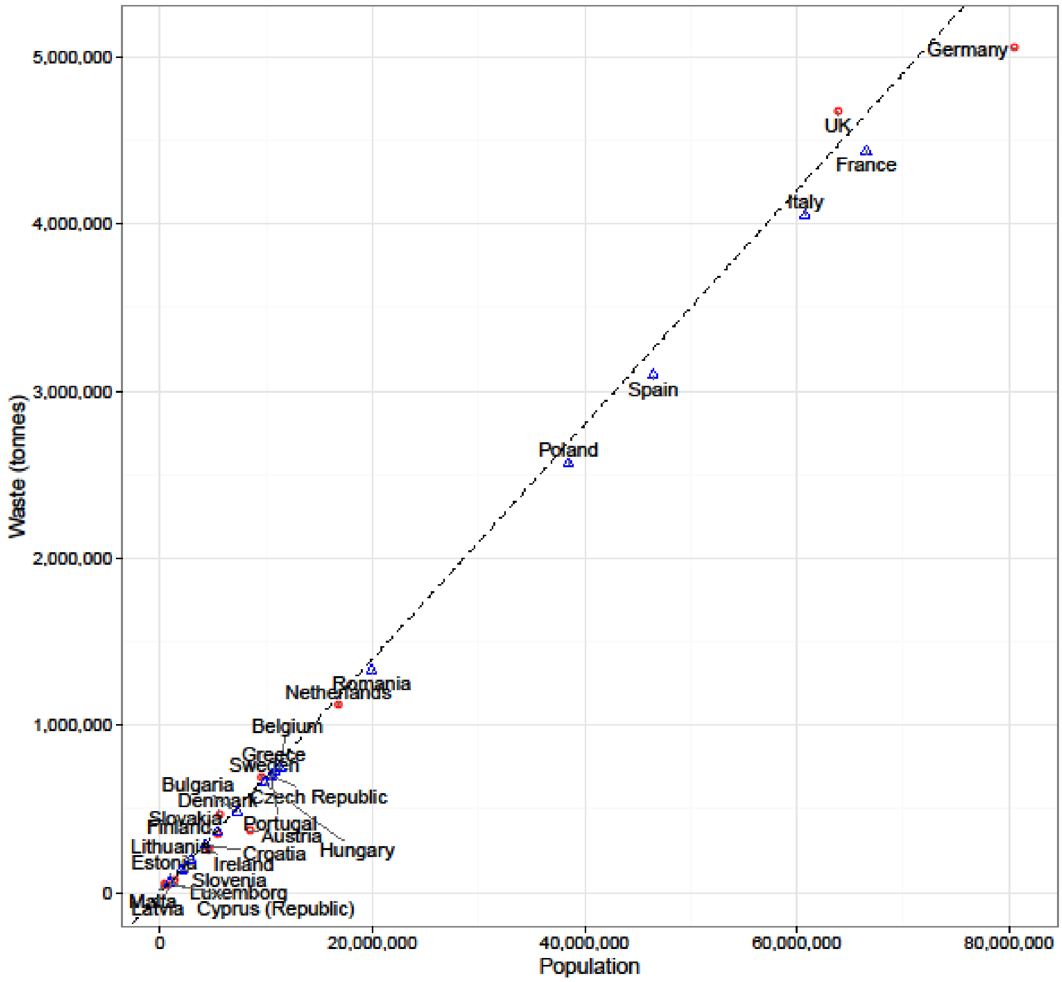


Figure 6. Municipal waste (tonnes) was strongly correlated to population size. Data from 11 countries (red symbols) were used to predict values for the remaining EU28 countries (blue symbols).

There was no apparent correlation between GDP and municipal waste per person (Figure 7). However, there was an apparent correlation between countries that spend greater than €3000 per capita on groceries and municipal waste. Greater levels of spend correlated with higher amounts of waste (Figure 8).

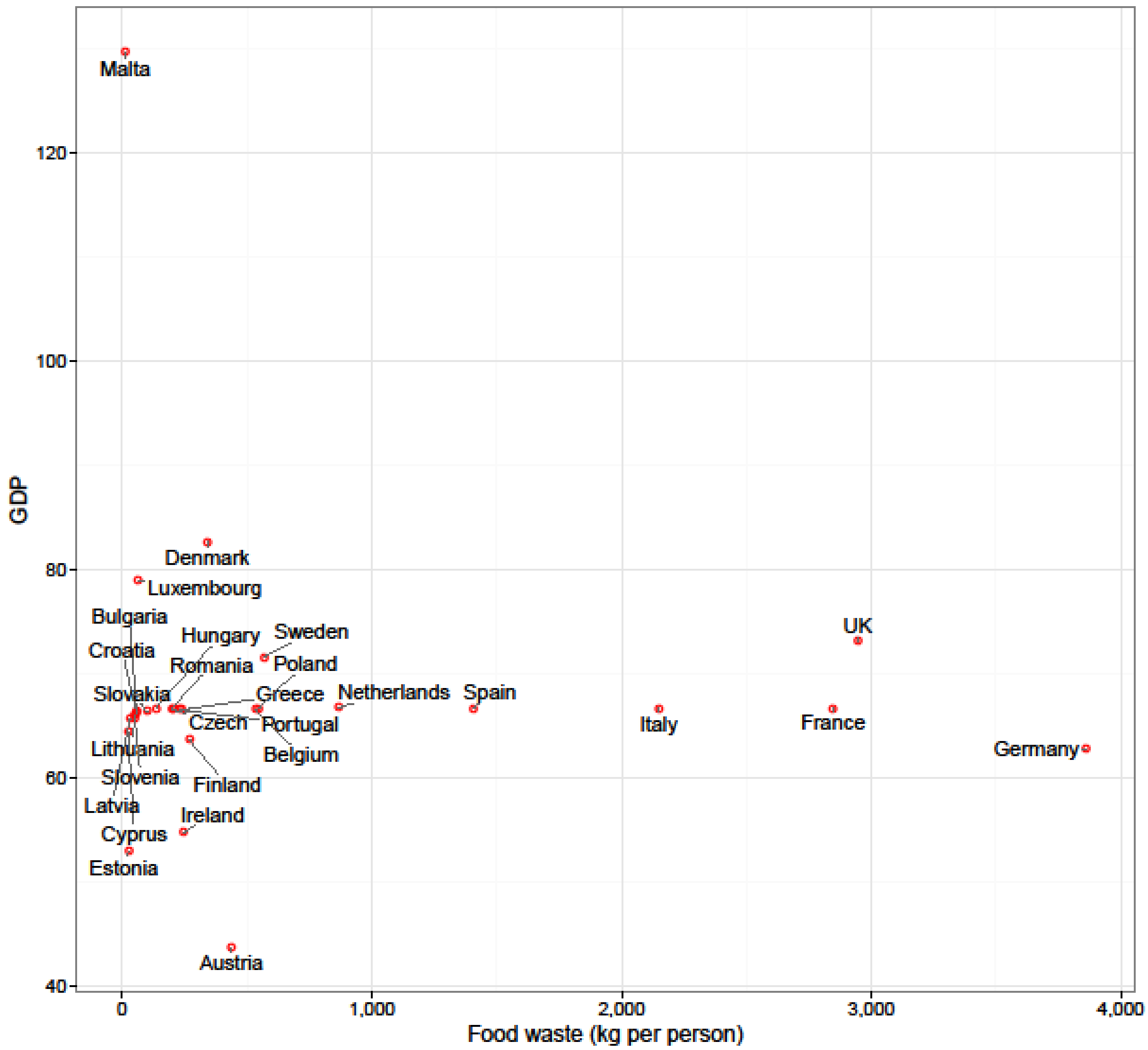


Figure 7. There was no apparent correlation between GDP and food waste per country.

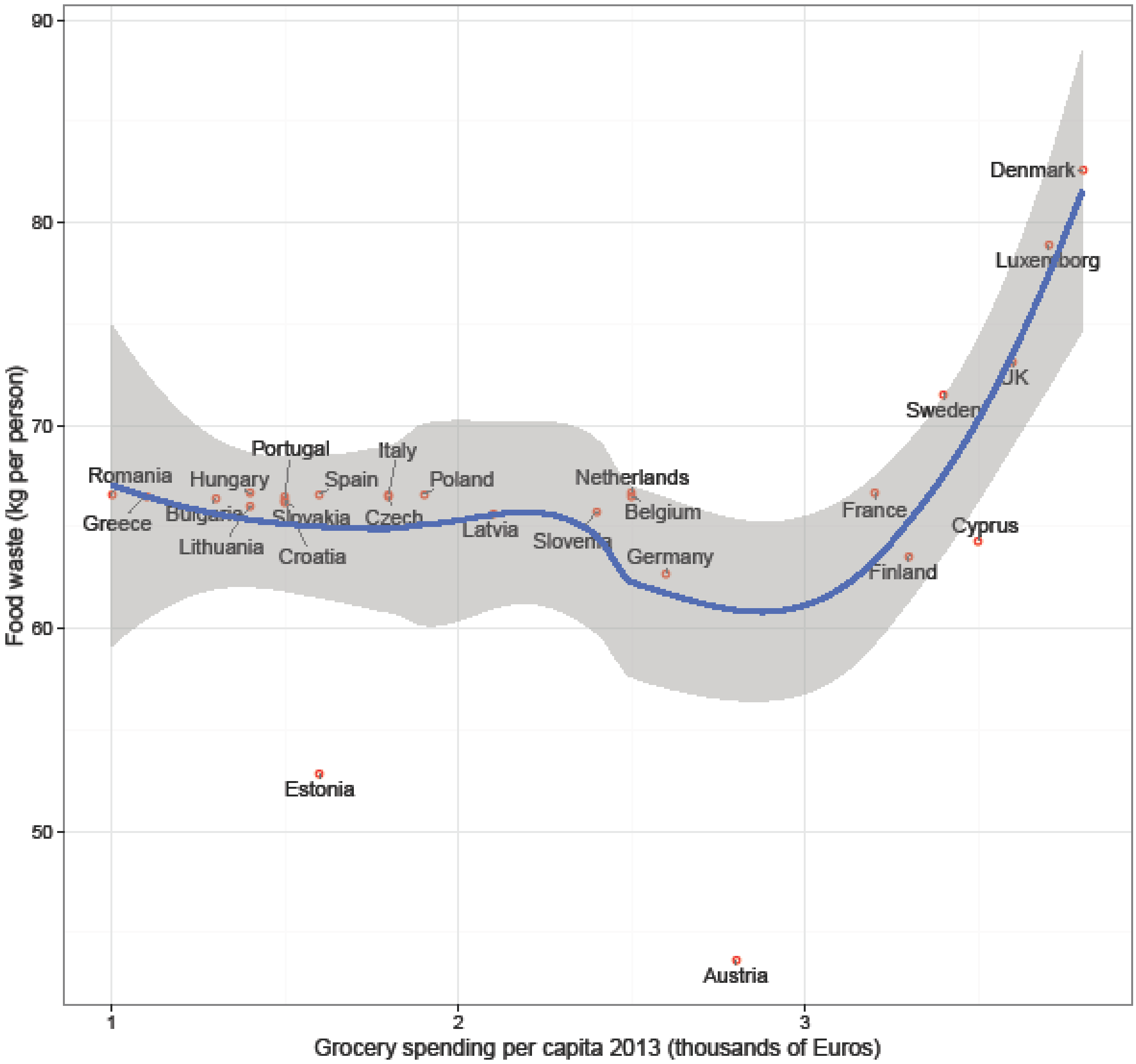


Figure 8. Grocery spend per capita (2013) was correlated to food waste but only at the highest levels of spend. The fitted loess regression line shown does come from a model with poor explanatory power.

Country-level differences in food waste generation means that data from one country cannot necessarily be applied to other countries. Therefore, the planned wide geographical assessment of food waste in the REFRESH project is validated.

3.4 Objective 2: Develop a systems map to illustrate potential links between consumer behaviour and the creation / reduction of food waste.

Bayesian networks are graphical models; hence they can be used to develop systems maps with defined probabilistic conditional relationships. Machine learning attempts to identify the most parsimonious structure using algorithms (such as the Hill-Climbing algorithm). In this process, however, some of the subtleties of the relationship between variables may be lost. From the structure of the Bayesian Network for dataset 1 (WRAP), for example, one can determine clear groupings (Figure 9). Discard behaviour, checking cupboards, shopping behaviours and social-demographic variables are clustered together. These groupings are consistent with the model of household food waste developed by REFRESH WP1 (D1_1) highlighting the importance of considering sequential waste generation across planning, retail and the home.

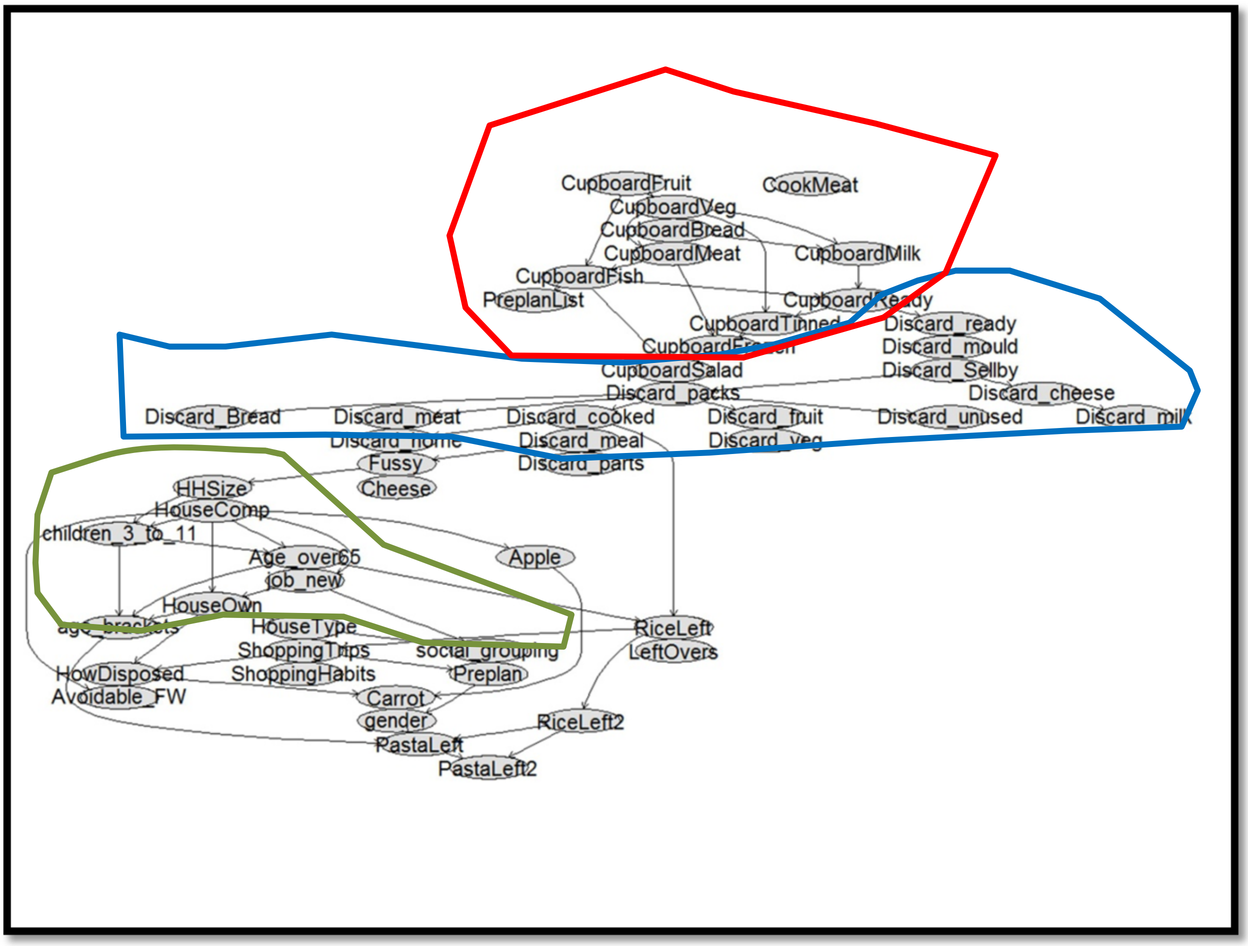


Figure 9. First iteration of a systems map developed as a machine-learnt Bayesian Network based on dataset 1 (WRAP) - UK consumers' food waste generation and responses to interviews on socio-economic status and behaviours. Groupings of behaviours (checking cupboards and discarding food) and social-demographic factors are indicated.

Bayesian networks can be structured or semi-structured prior to learning the relationships between variables. This way (as opposed to a machine learnt network) known, suspected or theoretical relationships between variables can be explored. In future iterations of the BN a structure can be imposed. This structure could reflect the consumer retail/food relationship (Figure 10). The acquisition of food takes several stages, first consumers may plan (by checking cupboards, writing a list, etc.), then consumers will acquire food at a shop(s), here behaviours related purchase become important (such as taking advantage of 2 for 1 offers, buying reduced cost items, etc.). Once food is purchased it is transported to the home and then stored, cooked or prepared and then consumed. Demographic variables (e.g. household size, composition, employment status, etc.) have some mediating effect on each of these stages of the process.

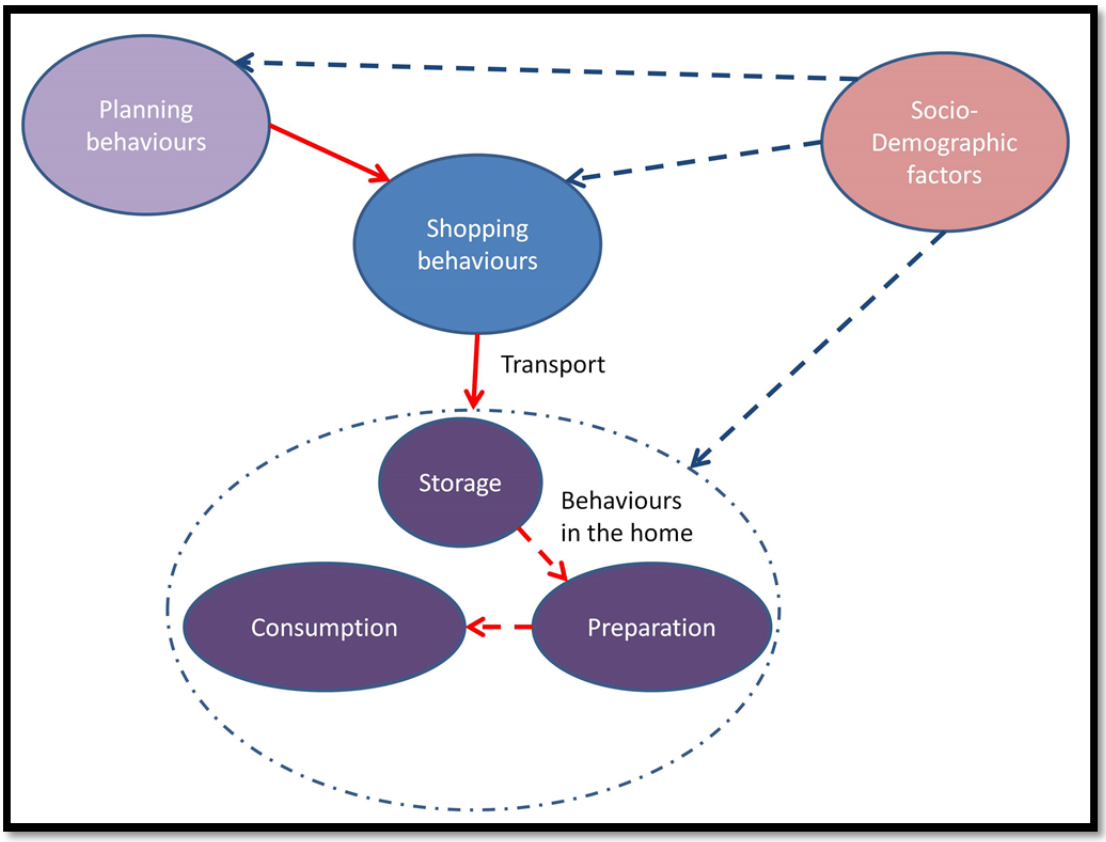


Figure 10. A conceptual structure for a BN model of the process of consumer food waste production

Box 1: Take outs

- People are inherently variable in how much waste they produce, therefore developing typologies is very important
- Household size and household composition are important drivers of food waste along with the demographic factors of age education and occupation
- Self-reporting food waste is not accurate
- There is a need for an objective (or semi-objective) measurement of food waste
- Country is an important driver of self-reported food waste
- Food waste increases with increased spending on groceries above a threshold of €3000 per annum
- There is a relationship between food waste and population at a country level but no relationship between food waste and GDP
- We need to understand the causal factors explaining country variation or we need to develop hierarchical models based on Country
- The conceptual classification outlined here has some validation from other Work Packages
- This structure should be incorporated in a semi-supervised way in future iterations of our analysis
- **A speculation** - The variation in the potential drivers of food waste is poorly understood. Household size is a consistently reported as a key driver of household food waste and as such it might be the best predictor from a policy perspective with additive adjustors for age, occupation, education, etc. Rather than exploring and understanding the full complexity (inclusive of all variation and uncertainty) of the relationships between demographics and food wasting behaviours one may be better placed to model a single variable (household size) instead.

4 Conclusions and future avenues for research

There is reasonable evidence that consumers are variable in their food waste behaviour i.e. food waste is not the same in different households. Household composition is an important driver of food waste, but it is not clear what the consequences of the interplay of planning, retail behaviour and in-home behaviour have on food waste. There is also evidence that consumers in different countries can be differentiated. There is some evidence to support age, education and occupation as typologies for consumer behaviour. Overall strength of evidence regarding consumer food waste is low, largely as a result of self-reporting underestimating food waste, unexplained inconsistency and low power.

Further work is required to ascertain why household composition is an important driver of food waste and the causes of country level variation in consumer food waste behaviour. Bayesian networks will be used to explore the former and hierarchical models will be used to explore the latter.

Planning, purchasing behaviour and behaviour in the home interact to produce domestic food waste but both interactions and structure are unclear. Machine-learning on new data, semi-supervised machine learning and expert elicited Bayesian networks based on the theory of planned behaviour will be used to explore these relationships further.

The high uncertainty regarding both structure and parameterisation of models indicate the need for a probabilistic approach to decision support if coherent evidence-informed policy is desired.

Box 2: Key findings

- Food waste is measured in many different ways, commonly weight is used but this may not be a useful comparator between waste types (calorific value may be a better unit of measurement for comparison across food types and countries, but may not be practical for field assessments). However weight may be necessary to perform life cycle analyses to assess greenhouse gas emissions or other parameters.
- Differences between geographical locations at different scales (country or local authority were important in both datasets) suggest that findings based exclusively on one country do not “translate” well to other contexts. Self-reporting waste may not give rise to accurate understanding of actual waste produced
- Avoidable food waste weight per household is a right-skewed distribution (Quested & Luzecka, 2014). The amount of waste produced might be a typology in itself.
- Typologies are difficult to elucidate at this stage. However, age, country, occupation, education, household size and composition all appear to be important variables.
- Discarding behaviour from the WRAP dataset is only a useful typology for those people who do not discard much waste.

5 References

Abeliotis, K. Lasaridi, K. & Chroni, C. (2014) Attitudes and behaviour of Greek households regarding food waste prevention. *Waste management & research*, **32** (2): 237 – 240.

Adelson, S.F., Delaney, I., Miller, C. & Noble, I.T. (1963) Discard of edible food n households. *Journal of home economics*, **55** (8): 633 – 638.

Bagherzadeh, M., Inamura, M. & Jeong, H. (2014) *Food Waste Along the Food Chain*. OECD Food, Agriculture and Fisheries Papers, No. 71, OECD Publishing. http://dx.doi.org/10.1787/5jxrcmftzj36-en

Baker, D., Fear, J. & Denniss, R. (2009) *What a waste: an analysis of household expenditure on food.* The Australia Institute, Canberra.

Barr, S. (2007) Factors influencing environmental attitudes and behaviours – A UK case study of household waste management. *Environment and Behaviour*, **39** (4): 435 – 473.

Bos-Brouwers, H.E.J., Soethoudt, J.M., Canali, M., Östergren, K. et al. (2012) *Drivers of current food waste generation, threats of future increase and opportunities for reduction*. EU Technical report (FUSIONS D1).

Canali, M., Ostergren, K., Amani, P., Aramyan, L. et al. (2013) *Report on food waste drivers for reducing food waste and barriers and opportunities*. EU Technical report (FUSIONS D1.5).

Flash Euro-barometer 388 (2013): Attitudes of Europeans towards waste management and resource efficiency (available at http://open-data.europa.eu/en/data/dataset/S1102_388)

George, R.M., Burgess, P.J. & Thorn, R.D. (2010) *Reducing food waste through the chill chain. Part 1: Insights around the domestic refrigerator*. WRAP, Banbury.

Glanz, R. (2008) *Causes of food waste generation in households an empirical analysis*. MSc thesis University of Natural Resources and Applied Life Sciences, Vienna and Cranfield University, UK.

IGD. (2007) *Beyond packaging*. http://www.fcrn.org.uk/research-library/igd-report-beyond-packaging

Johnson, D., Hipps, N. & Hails, S. (2009) *Helping consumers reduce fruit and vegetable waste*. WRAP, Banbury.

Katajajuuri et al. (2012) Katajajuuri, J., Silvennoinen, K., Hartikainen, H., Jalkanen, L., Koivupuro, H. & Reinikainen, A. (2012) *Food waste in the food chain and related climate impacts.* 8th International Conference on Life Cycle Analysis, 1-4 October, Saint- Malo, France.

Koivupuro, H-K., Hartikainen, H., Silvennoinen, K., Katajajuuri, J-M., Heikintalo, N., Reinikainen, A. & Jalkanen, L. (2012) Influence of socio-demographical, behavioural and attitudinal factors on the amount of avoidable food waste generated in Finnish households. *International Journal of Consumer Studies*, **36**: 183 – 191.

Møller, H., Hanssen, O.J., Gustavsson, J., Östergren, K. et al. (2013) *Report on review of (food) waste reporting methodology and practice*. EU Technical report (FUSIONS D1.3).

Muth, M.K., Kosa, K.M., Nielsen, S.J. & Karns, S.A. (2007) *Exploratory research on estimation of consumer-level food loss conversion factors*. Final report, RTI International. Research Triangle Park (NC - USA).

Parfitt, J., Barthel, M. and Macnaughton, S. (2010). Food waste within food supply chains: quantification and potential for change to 2050. *Philosophical Transactions of the Royal Society B – Biological Sciences*, **365** (1554): 3065 – 3081.

Parizeau, K. von Massow, M. & Martin, R. (2015). Household-level dynamics of food waste production and related beliefs, attitudes and behaviours in Guelph, Ontario. *Waste Management*, **35**: 207 – 217.

Pekcan, G., Köksal, E., Kücükerdonmez, Ö. & Özel, H. (2005). *Household food waste in Turkey*. Food and Agriculture Organisation of the United Nations, Statistics Division Working paper series ESS/ESSA/006e.

Quested, T.E. & Luzecka, P. (2014) Household food and drink waste: A people focus. WRAP, Banbury, UK.

Quested, T.E., Marsh, E., Stunell, D. & Parry, A.D. (2013) Spaghetti soup: The complex world of food waste behaviours. *Resources, Conservation and Recycling*, **79**: 43 – 51.

Quested, T.E., Parry, A.D., Easteal, S. & Swannell, R. (2011) Food and drink waste from households in the UK. *Nutrition Bulletin*, **36**: 460 – 467.

Secondi, L., Principato, L. & Laureti, T. (2015) Household food waste behaviour in EU-27 countries: A multilevel analysis. *Food Policy*, **56**: 25 – 40.

Sibrián, R., Komorowska, J. & Mernies, J. (2016) Estimating household and institutional food wastage and losses. . Food and Agriculture Organisation of the United Nations, Statistics Division Working paper series ESS/ESSA/001e.

Silvennoinen, K., Katajajuuri J.M., Hartikainen, H., Koivupuro, H.K., Jalkanen L. & Reinikainen A. (2012) *Food waste volume and composition in the Finnish supply chain: special focus on food service sector*. Proceedings of the Fourth International Symposium on energy from biomass and waste, Venice, Italy, 12 – 15 November 2012.

Sonesson, U., Anteson, F., Davis, J. & Sjödén. (2005) Home transport and wastage: Environmentally relevant household activities in the life cycle of food. *Ambio,* **34** (4/5): 371 – 375.

Stancu, V., Haugaard, P., Lähteenmäki, L. (2016) Determinants of consumer food waste behaviour: Two routes to food waste. *Appetite*, **96**: 7 – 17.

Stefan, V., van Herpen, E., Tudoran, A.A. & Lähteenmäki, L. (2013) Avoiding food waste by Romanian consumers: The importance of planning and shopping routines. *Food Quality and Preference,* **28**: 375 – 381.

Stenmarck, Å., Jensen, C., Quested, T.E. & Moates, G. (2016) Estimates of European food waste levels. FUSIONS report http://www.eu-
fusi-
si-
ons.org/phocadownload/Publications/Estimates%20of%20European%20food%20
waste%20levels.pdf

van Garde S.J. & Woodburn, M.J. (1987) Food discard practices of householders. Journal of the American Dietary Association, **87**: 322 – 329.

Wasserman, G. & Schneider, F. (2005) *Edibles in household waste*. Proceedings of the Tenth International Waste Management and Landfill Symposium, S. Margherita di Pula, Cagliari, Italy; 3 - 7 October 2005

Wenlock, R.H. & Buss, D.H. (1977) Wastage of edible food in the home: a preliminary study. *Journal of Human Nutrition*, **31** (6): 405 – 411.

Wenlock, R.H., Buss, D.H., Derry, B.J. & Dixon, E.J. (1980) Household food wastage in Britain. *British Journal of Nutrition*, **43** (1): 53 – 70.

WRAP & French-Brooks, J. (2009) *Reducing supply chain and consumer potato waste*. WRAP, Banbury, UK.

WRAP (2007) *We don't waste food! A householder survey.* WRAP (2008), Banbury, UK.

WRAP (2008) The food we waste. WRAP, Banbury, UK.

WRAP (2009) *Down the drain*. WRAP, Banbury, UK.

WRAP (2011) *Investigation into the possible impact of promotions on food waste*. WRAP, Banbury, UK.

WRAP (2012) *Household food and drink waste in the United Kingdom 2012*. WRAP, Banbury, UK.

6 Annexes

6.1 Technical appendix

The following sections outline the more technical aspects of the methodological approach and the results.

6.2 Methods for the Regression analysis

Dataset 1 (WRAP's Waste compositional analysis)

Generalised Linear Models (GLMs) were applied to WRAP's "Waste compositional analysis" to assess correlations between "avoidable waste" and a number of socio-demographic and behavioural variables (Table A1). All categorical variables were treated as factors in the analysis. This means that data such as age which were recorded in three categories were not treated as continuous dummy variables (e.g. 1,2,3) which can violate the assumptions of the model used.

Table A1. The variables used in the development of regression models assessing the drivers of consumer food waste[3].

Variables
Food waste
Avoidable waste; defined as "food and drink thrown away that was, at some point prior to disposal, edible, e.g. milk, lettuce, fruit juice, meat (excluding bones, skin, etc.)"
Socio-demographic characteristics of the respondent
Age
Gender
Socio-demographic characteristics of the household
Age structure, based on ages of all household members
Household size (i.e. number of people in the household)
Location (rural or urban area (kitchen diary only))
Household composition, e.g. a couple, or family with young children
Home ownership status, e.g. privately rented or owned with mortgage
Type of residence, e.g. flat or terraced house
Presence of children
Social-economic status – calculated based on the characteristics of the main earner
Socio-demographic characteristics of the main earner
Employment status
Consumer behaviour
The extent of meal planning
Cupboard checking
List making
Use of freezer
Storage of cheese and meats after opening
Use of the fridge to store apples and carrots
Using leftovers
Cooking the right amount of rice and pasta

Throwing away items because they have gone past their date label
Type of shopping trips made
Frequency of main shopping trip
Proportion of occupants of the household classed by the survey respondent to be fussy eaters

[3] The definitions of "avoidable food waste" and demographic and social variables are those used in Quested & Luzecka (2014), these may differ to those of the FUSIONS project and in other assessments of consumer food waste.

Highly co-linear variables (those with a square-root of variance inflation factor greater than 2) were removed. The potential number of models was greater than 1 billion (36 predicator variables remained in the model after removing highly co-linear variables meaning that the potential number of combinations was 2^n or 68719476736) so the package random forest (Liaw & Wiener 2002; see further description below) in the R programme (R Core Team 2016) was used to identify the most important (see below) variables. Then AICc was used to determine a set of plausible models and modelling averaging used to obtain estimates of the effect of predictors on avoidable food waste (AIC is a method by which the relative quality of statistical models for a given set of data can be estimated, it allows the identification of a plausible set of models which can then be averaged).

Dataset 2 (Flash Euro-barometer 388: Attitudes of Europeans towards waste management and resource efficiency)

GLM's were also applied to the Flash Euro-barometer 388 dataset using the self-reported percentage of food purchased that goes to waste. As the response variable was categorical multinomial models were used. The predictors included the variables listed in Table A2.

Table A2. The variables used in the development of regression models assessing the drivers of consumer food waste from Euro-barometer 388.

Variables
Food waste
Estimate of the percentage of the food purchased that goes to waste
Socio-demographic characteristics of the respondent
Age
Gender
Nationality
Age full-time education stopped
Current occupation (self-employed, an employee, a manual worker or without a professional activity)
Socio-demographic characteristics of the household
Location (rural/urban/etc.)
Phone ownership
Household composition (aged 15 and over only)
Attitudes to waste and resource use
How important is it that Europe uses its resources efficiently
What would the impact of efficient resource use be on Economic growth
What would the impact of efficient resource use be on Employment
What would the impact of efficient resource use be on Quality of life

Which policies would make the biggest difference in resource use efficiency (Cutting taxes on employment and increasing taxes on resource use, reducing waste and sorting recyclable waste at home, reducing and recycling waste in industry and construction)

How much do you agree with the statement "[OUR COUNTRY] as a whole is generating too much waste"

How much do you agree with the statement "Your household is generating too much waste"

How much do you agree with the statement "You make efforts to reduce the amount of household waste that you generate"

What actions do you take to reduce household waste

Do you sort waste (paper, plastic, metal, glass, kitchen waste, garden waste, hazardous waste, electrical products)

What would convince you to sort your waste (More convenient separate waste collection at your home, More and better waste recycling and composting facilities in your area, More information on how and where to separate waste, Increased tariffs if waste is not separated properly, and Financial incentives to separate waste (deposits, reduced tariffs, etc.)

How would you manage the cost of household waste (through a flat rate, a contribution relative to your waste production, or more producer responsibility) which would be preferable

What would help you to waste less food (Better and clearer information on how to interpret 'best before' dates, Better and clearer information on food product labels, e.g. information on storage and preparation, Better shopping planning by your household, Better estimation of portion sizes (how much food you cook) to avoid wasting food, Availability of smaller portion sizes in shops, Re-using leftovers instead of throwing them away)

What aspects of durable/electrical goods do you consider when making a purchase (You can use the product for a long time, The producer gives you a longer warranty/guarantee for the product, The product is made from recycled materials, The product can be recycled after you use it, The product is environmentally-friendly, You can easily sell the product when you no longer want to use it, The seller will take back the old product when you buy a new one, The running costs are lower due to greater efficiency)

Would you buy; Textiles, Electronic equipment, Furniture, Household electrical appliances, Books, CDs, DVDs, video games, second hand?

Q14 There are emerging alternatives to buying new products. Have you ever done any of the following? (Bought a remanufactured product. This is a used product, the faulty or old components of which have been substituted, and which is sold with the same guarantees as a new product, Leased or rented a product instead of buying it (e.g. a washing machine, furniture), Used sharing schemes. These can be organised, like car or bike sharing schemes, or informal, like neighbours sharing lawn mowers)

What are your opinions on how to address the problems of plastic waste

How much litter is there in the area where you live?

What would be the most efficient in reducing littering? (Organised clean-up events, Better enforcement of existing anti-litter laws, Encouraging alternatives to plastic bags or other plastic packaging, Increasing and encouraging the recycling of waste, Communication campaigns to raise awareness among citizens, Ensuring availability of public litter bins, Financial participation by producers of plastics in funding the fight against litter)

The amount of litter entering the oceans is a cause for concern. Would you support the development of an EU-level target to reduce such litter?

Once again, highly co-linear variables were removed and Random Forests used to further reduce the variable set. AICc was used to determine a set of plausible models and model averaging used.

6.3 Methods for the Machine learning (Random forest & Bayesian network)

Dataset 1 (WRAP)

We used random forests (Breiman 2001) on the complete dataset to ascertain if the same drivers of food waste were identified. Random forests is an ensemble learning (i.e. machine learning) algorithm for classification and regression. In essence the algorithm constructs multiple (minimum of 500) “regression” or “classification” trees (also known as “decision trees”) and outputs the mode of the classes (or mean predictions). The advantage of this method above classification and regression trees is that it is robust to multicollinearity and corrects for over fitting (i.e. having a model that cannot be generalised to novel data). See Breiman (2001) for more information.

We also used a machine learnt Bayesian network (using the hill-climbing algorithm in the bnlearn package in R; Scutari & Denis 2014) predicting total avoidable food waste to develop a candidate causal network of avoidable food waste. For speed of processing we discretised the dataset (to convert continuous data to categorical data, using the quartiles to determine the categories) and used this to build the models.

Dataset 2 (Flash Euro-barometer 388)

The Flash Euro-barometer data were machine learnt as above using both Random Forests and a Bayesian network.

6.4 Results for the Regression analysis

Dataset 1 (WRAP)

There were high levels of multi co-linearity in the dataset with some variables having perfect correlation. After these were removed we used random forests to further reduce the model set to 30 variables (the largest number of variables that R memory allocation can process for model selection and averaging – this equates to an approximate 1073741824 models).

A random forest identified discard behaviours (relating to vegetables, bread, homemade meals and inedible parts), household size, household composition and local authority as important variables for the regression model. In the final model set (the most parsimonious models given the data and variables selected) discarding behaviour relating to bread and vegetables, household size and local authority were in all three of the top models (Table A3).

Variables with the largest effect (and smallest standard error) on avoidable food waste included discard behaviour in relation to vegetables, with those people responding to the question of how much they had thrown away in the last week. Discarding “none”, “hardly any” or a “small amount” of vegetables was related to a decrease avoidable food waste. Similarly discarding “none”, “hardly any”, or a “small amount” of inedible parts of food was related to a decrease in avoidable waste. Households size (5 or 6 and over people) and discarding “some” bread was related to high avoidable food waste (Table A3).

Table A3. Model averaging of the most parsimonious combined model set for dataset 1 (WRAP)

	Variable	Estimate	Std.Error	Adjusted SE	z value
	(Intercept)	2407.65	784.72	785.23	3.07
Discard behaviour	Bread: A small amount	308.18	354.66	354.91	0.87
Discard behaviour	Bread: Don't eat it	−243.98	482.77	483.11	0.51
Discard behaviour	Bread: Hardly any	209.57	358.87	359.12	0.58
Discard behaviour	Bread: None	−28.02	346.96	347.21	0.08
Discard behaviour	Bread: Quite a lot	−627.46	762.73	763.27	0.82
Discard behaviour	Bread: Some	523.12	385.69	385.96	1.36
Discard behaviour	Inedible parts: A small amount	−224.69	177.73	177.78	1.26
Discard behaviour	Inedible parts: Don't eat it	−338.67	427.73	427.96	0.79
Discard behaviour	Inedible parts: Hardly any	−284.47	223.83	223.89	1.27
Discard behaviour	Inedible parts: None	−357.66	264.44	264.49	1.35
Discard behaviour	Inedible parts: Quite a lot	5.22	149.33	149.43	0.04
Discard behaviour	Inedible parts: Some	−200.25	172.46	172.51	1.16
Discard behaviour	Vegetables: A small amount	−1394.56	411.92	412.21	3.38
Discard behaviour	Vegetables: Don't eat it	−1331.62	523.37	523.73	2.54
Discard behaviour	Vegetables: Hardly any	−1466.38	415.01	415.30	3.53
Discard behaviour	Vegetables: None	−1465.49	404.49	404.77	3.62
Discard behaviour	Vegetables: Quite a lot	−1022.36	830.04	830.63	1.23
Discard behaviour	Vegetables: Some	−1032.59	436.29	436.59	2.37
Household size	2	485.28	567.47	567.81	0.86
Household size	3	1275.56	541.71	542.07	2.35
Household size	4	1445.60	535.77	536.13	2.70
Household size	5	1802.28	553.23	553.61	3.26
Household size	6+	1766.16	565.53	565.92	3.12
Local authority	Blaenau Gwent	9.22	226.09	0.04	0.97
Local authority	Bridgend	−381.41	193.96	194.09	1.97
Local authority	Cannock Chase	−318.08	219.05	1.45	0.15
Local authority	Milton Keynes	−114.94	210.56	0.55	0.59
Local authority	Neath Port Talbot	19.32	0.10	0.92	
Local authority	Poole	−34.25	188.04	188.18	0.18
Local authority	Scarborough	−436.31	205.08	205.23	2.13
Local authority	Slough	730.77	236.05	236.22	3.09
Local authority	Suffolk Coastal	−354.28	206.00	1.72	0.09
Local authority	Warrington	−635.79	256.39	256.57	2.48
Local authority	Wyre Forest	−167.73	207.48	0.81	0.42
Household composition	Family with at least one child under 18	83.20	176.01	176.05	0.47
Household composition	Family with child(ren) (all 18 years or over)	20.00	101.07	101.13	0.20
Household composition	Other Not stated	−41.07	160.37	160.46	0.26

Household composition	Single occupancy	136.31	567.15	567.48	0.24

Dataset 2 (Flash Euro-barometer 388)

A random forest identified country, age, household size, employment opportunities, house management, litter, economic growth, self-reporting your family wastes too much, community type, education and occupation as important variables for the regression model. In the final model set (the most parsimonious models given the data and variables selected) all these variables were retained.

Variables with the largest effect (and smallest standard error) on self-reported food waste included age, house management, country, education, self-reporting that your family wastes too much and household size.

Table A4. Model averaging of the most parsimonious combined model set for dataset 2 (Euro-barometer 388).

	None			5pc_or_less			6pc_to_15pc			31pc_to_50pc			More_than_50pc			DKNA		
	Coefficient	Error	Z-score	Coefficient	Error	Z-score	Coefficient	Error	Z-score	Coefficient	Error	Z-score	Coefficient	Error	Z-score	Coefficient	Error	Z-score
(Intercept)	-1.89	0.72	-2.64	0.76	0.62	1.22	-0.68	0.67	-1.02	-0.96	1.70	-0.57	-2.03	1.87	-1.09	-1.05	0.91	-1.15
CountryBelgium	-0.43	0.29	-1.48	0.07	0.21	0.34	0.76	0.23	3.34	-0.78	0.53	-1.47	-0.21	0.72	-0.28	0.56	0.54	1.05
CountryBulgaria	0.92	0.26	3.51	0.09	0.22	0.41	0.20	0.24	0.86	-0.39	0.49	-0.79	0.66	0.63	1.04	1.06	0.51	2.10
CountryCroatia	-0.56	0.26	-2.17	-0.60	0.20	-3.03	-0.02	0.21	-0.10	-1.00	0.49	-2.05	-0.31	0.67	-0.47	-0.69	0.62	-1.11
CountryCyprus_(Republic)	-0.17	0.30	-0.56	-0.64	0.23	-2.81	-0.28	0.25	-1.12	0.17	0.45	0.39	0.17	0.70	0.24	1.28	0.51	2.52
CountryCzech_Republic	1.21	0.32	3.79	1.19	0.27	4.41	0.82	0.29	2.85	-0.48	0.67	-0.72	0.12	0.88	0.14	1.40	0.58	2.42
CountryDenmark	-2.04	0.36	-5.66	-0.44	0.20	-2.16	0.59	0.22	2.69	-0.38	0.47	-0.81	-0.01	0.68	-0.01	0.62	0.50	1.24
CountryEstonia	1.95	0.30	6.43	0.84	0.27	3.09	0.81	0.29	2.81	-1.37	0.94	-1.46	-1.37	1.58	-0.86	1.30	0.56	2.34
CountryFinland	-0.26	0.28	-0.93	-0.09	0.22	-0.43	0.66	0.23	2.87	-0.77	0.56	-1.38	-1.13	1.02	-1.11	0.25	0.56	0.44
CountryFrance	-0.08	0.32	-0.24	0.72	0.24	3.00	0.97	0.25	3.82	-0.67	0.60	-1.12	1.01	0.66	1.54	1.27	0.54	2.36
CountryGermany	-1.22	0.32	-3.82	-0.03	0.21	-0.12	0.37	0.23	1.62	-0.10	0.46	-0.21	0.34	0.69	0.49	-0.38	0.63	-0.61
CountryGreece	-0.59	0.27	-2.20	-0.69	0.19	-3.54	-0.12	0.21	-0.57	-0.32	0.42	-0.77	0.15	0.62	0.24	1.24	0.47	2.63
CountryHungary	0.92	0.26	3.55	-0.01	0.22	-0.02	0.36	0.24	1.53	-1.12	0.60	-1.88	0.43	0.65	0.66	0.94	0.51	1.82
CountryIreland	-1.89	0.33	-5.78	-0.54	0.20	-2.70	0.22	0.22	1.00	-0.10	0.43	-0.23	-0.38	0.72	-0.52	0.10	0.53	0.19
CountryItaly	-0.45	0.27	-1.65	-0.35	0.20	-1.74	0.26	0.22	1.18	-0.90	0.49	-1.83	-0.15	0.68	-0.22	1.39	0.47	2.92
CountryLatvia	0.75	0.25	2.97	-0.32	0.21	-1.52	0.12	0.23	0.51	0.34	0.41	0.84	0.70	0.62	1.11	1.32	0.48	2.75

CountryLithuania	1.55	0.25	6.28	0.04	0.21	0.18	-0.09	0.23	-0.39	-0.77	0.51	-1.52	-0.60	0.77	-0.79	0.17	0.55	0.30
CountryLuxembourg	-2.53	0.73	-3.47	0.26	0.28	0.94	0.77	0.29	2.62	0.25	0.55	0.45	0.15	0.88	0.18	-0.87	1.06	-0.82
CountryMalta	2.18	0.50	4.35	1.55	0.47	3.25	1.63	0.49	3.33	0.16	0.89	0.19	-0.84	2.22	-0.38	2.58	0.68	3.79
CountryPoland	1.01	0.26	3.89	0.14	0.21	0.65	0.05	0.23	0.19	-1.43	0.66	-2.18	-0.88	0.86	-1.02	-0.02	0.60	-0.04
CountryPortugal	0.65	0.27	2.43	0.03	0.22	0.14	0.52	0.23	2.21	0.31	0.42	0.74	0.87	0.61	1.42	2.61	0.46	5.63
CountryRomania	1.35	0.25	5.32	0.11	0.21	0.51	0.37	0.23	1.62	0.06	0.43	0.14	0.31	0.63	0.49	0.25	0.57	0.44
CountrySlovakia	1.24	0.26	4.76	0.20	0.22	0.92	0.24	0.24	1.01	-0.19	0.46	-0.42	-0.31	0.77	-0.41	1.50	0.49	3.04
CountrySlovenia	0.78	0.28	2.75	0.53	0.24	2.21	0.74	0.26	2.86	0.73	0.44	1.66	0.81	0.66	1.24	0.07	0.66	0.10
CountrySpain	0.94	0.29	3.27	0.83	0.23	3.59	0.89	0.25	3.57	-0.49	0.54	-0.90	0.08	0.73	0.11	0.43	0.63	0.68
CountrySweden	-1.31	0.29	-4.45	-0.52	0.20	-2.61	0.37	0.21	1.73	-3.65	1.54	-2.36	-2.61	1.57	-1.66	0.40	0.51	0.78
CountryThe_Netherlands	-1.50	0.32	-4.75	-0.35	0.20	-1.72	0.33	0.22	1.53	-1.27	0.59	-2.14	-0.79	0.86	-0.92	-1.04	0.72	-1.44
CountryUnited_Kingdom	-1.06	0.29	-3.70	-0.39	0.20	-1.91	0.35	0.22	1.58	-0.93	0.51	-1.83	0.07	0.67	0.11	0.16	0.53	0.30
Age25 -34	1.27	0.25	5.16	0.72	0.12	5.84	0.40	0.13	3.19	-0.24	0.25	-0.96	-0.13	0.33	-0.41	0.14	0.29	0.48
Age35 - 44	2.16	0.24	8.95	1.21	0.13	9.58	0.65	0.13	4.95	-0.54	0.27	-1.96	-0.17	0.35	-0.49	0.51	0.28	1.81
Age45 - 54	2.75	0.24	11.61	1.54	0.12	12.37	0.57	0.13	4.41	-0.52	0.27	-1.92	-0.41	0.36	-1.13	0.87	0.27	3.17
Age55 - 64	3.26	0.24	13.70	1.92	0.13	14.68	0.79	0.14	5.75	-0.43	0.29	-1.51	0.08	0.35	0.22	1.29	0.27	4.73
Age65 & older	3.66	0.24	15.09	2.09	0.14	15.06	0.65	0.15	4.43	-0.15	0.30	-0.50	-0.17	0.39	-0.42	1.43	0.28	5.11
Household_size.152	-0.02	0.10	-0.18	0.11	0.08	1.40	0.22	0.09	2.54	0.14	0.20	0.69	0.11	0.26	0.43	-0.12	0.16	-0.78
Household_size.153	-0.08	0.12	-0.70	0.04	0.10	0.45	0.11	0.10	1.09	0.11	0.23	0.47	-0.10	0.32	-0.32	0.02	0.19	0.10
Household_size.154+	0.18	0.13	1.42	0.22	0.10	2.17	0.33	0.11	3.00	0.08	0.24	0.34	0.50	0.30	1.65	0.22	0.20	1.11
Household_size.15DK	0.44	2.60	0.17	1.13	2.37	0.47	-2.40	4.13	-0.58	-0.23	6.58	-0.03	-0.18	7.98	-0.02	2.71	2.48	1.09
Household_size.15Refusal	0.26	0.63	0.42	0.06	0.54	0.12	0.06	0.58	0.11	0.20	1.17	0.17	0.90	1.20	0.75	0.77	0.75	1.02
Employment_opportunitiesSomewhat_negative	-0.29	0.18	-1.56	0.00	0.15	-0.01	0.09	0.17	0.56	0.43	0.40	1.07	0.25	0.48	0.53	-0.42	0.28	-1.53
Employment_opportunitiesSomewhat_positive	-0.27	0.15	-1.74	0.15	0.13	1.12	0.28	0.14	2.00	0.34	0.36	0.95	-0.23	0.43	-0.52	-0.39	0.22	-1.76
Employment_opportunitiesVery_negative	-0.52	0.22	-2.43	-0.10	0.18	-0.53	-0.10	0.20	-0.49	0.51	0.43	1.19	0.81	0.48	1.67	-0.48	0.31	-1.55
Employment_opportunitiesVery_positive	-0.11	0.16	-0.65	0.20	0.14	1.46	0.18	0.15	1.20	0.29	0.37	0.79	0.04	0.44	0.09	-0.53	0.23	-2.26
House_managementDK/NA	1.17	0.16	7.47	0.69	0.14	4.80	0.31	0.15	2.04	0.59	0.28	2.09	0.44	0.37	1.18	1.29	0.21	6.16
House_managementFixed sum	0.39	0.11	3.67	0.23	0.08	2.79	0.24	0.09	2.80	0.16	0.18	0.89	0.21	0.24	0.85	0.66	0.17	3.97
House_managementPay in proportion	0.39	0.09	4.48	0.31	0.07	4.56	0.25	0.07	3.53	-0.18	0.16	-1.07	-0.01	0.21	-0.03	0.30	0.15	2.01
LitterDKNA	0.25	0.48	0.53	0.29	0.43	0.66	0.56	0.46	1.24	-0.54	1.10	-0.50	0.00	1.09	0.00	1.59	0.50	3.17

LitterNone	0.46	0.14	3.43	0.09	0.11	0.82	-0.05	0.12	-0.45	-0.23	0.25	-0.91	-0.53	0.31	-1.71	0.23	0.20	1.11
LitterNot_much	0.02	0.12	0.20	0.17	0.09	1.87	0.31	0.10	3.11	-0.30	0.21	-1.47	-0.43	0.24	-1.80	-0.05	0.18	-0.29
LitterQuite_a_lot	-0.09	0.13	-0.72	0.07	0.10	0.65	0.26	0.11	2.45	-0.16	0.22	-0.74	-0.61	0.28	-2.21	-0.14	0.19	-0.71
Economic_growthSomewhat_negative	-0.55	0.20	-2.78	-0.32	0.17	-1.89	-0.11	0.18	-0.60	-0.17	0.38	-0.45	-0.62	0.50	-1.24	-0.44	0.30	-1.49
Economic_growthSomewhat_positive	-0.30	0.17	-1.78	-0.05	0.15	-0.33	0.15	0.16	0.91	-0.22	0.34	-0.64	-0.18	0.43	-0.42	-0.32	0.25	-1.31
Economic_growthVery_negative	-0.35	0.24	-1.47	-0.23	0.20	-1.10	-0.37	0.23	-1.62	0.07	0.44	0.16	-0.10	0.52	-0.20	-0.25	0.35	-0.73
Economic_growthVery_positive	-0.19	0.18	-1.07	0.01	0.16	0.04	0.13	0.17	0.79	-0.12	0.36	-0.33	-0.30	0.45	-0.67	-0.20	0.26	-0.78
Too_much_wasteTend_to_agree	-0.92	0.35	-2.62	-0.16	0.32	-0.52	0.57	0.37	1.57	-0.84	0.56	-1.50	-0.33	0.95	-0.34	-1.37	0.41	-3.37
Too_much_wasteTend_to_disagree	-0.17	0.35	-0.48	0.50	0.32	1.56	0.87	0.37	2.37	-0.81	0.56	-1.43	-0.46	0.96	-0.48	-0.89	0.41	-2.19
Too_much_wasteTotally_agree	-1.06	0.35	-3.00	-0.47	0.32	-1.48	0.35	0.37	0.96	-0.72	0.56	-1.29	0.22	0.95	0.23	-1.21	0.41	-2.95
Too_much_wasteTotally_disagree	0.27	0.35	0.76	0.67	0.32	2.08	0.63	0.37	1.72	-0.50	0.57	-0.89	0.57	0.95	0.61	-0.63	0.41	-1.55
Community_typeLarge_town	-0.42	0.52	-0.81	-0.24	0.47	-0.51	-0.12	0.49	-0.25	0.57	1.50	0.38	0.37	1.43	0.26	-0.54	0.60	-0.90
Community_typeRural_area_or_village	0.04	0.52	0.08	-0.04	0.47	-0.09	-0.24	0.49	-0.49	0.89	1.50	0.59	0.42	1.43	0.29	-0.42	0.60	-0.69
Communi-ty_typeSmall_or_middle_sized_town	-0.33	0.52	-0.63	-0.19	0.46	-0.41	-0.16	0.49	-0.32	1.09	1.50	0.73	0.30	1.43	0.21	-0.48	0.60	-0.81
Education20_years_and_older	-0.52	0.09	-6.07	-0.26	0.07	-3.69	0.06	0.07	0.76	-0.32	0.17	-1.94	-0.19	0.22	-0.87	-0.26	0.14	-1.86
EducationDK	0.17	0.37	0.47	-0.11	0.34	-0.32	-0.02	0.37	-0.05	0.12	0.68	0.18	0.96	0.69	1.39	1.08	0.43	2.53
EducationNo_full-time_education	0.96	0.47	2.04	0.01	0.42	0.02	-0.25	0.46	-0.55	-2.09	2.27	-0.92	1.38	0.75	1.85	0.63	0.66	0.95
EducationRefusal	0.74	1.07	0.69	0.30	1.01	0.29	-0.35	1.11	-0.32	-0.76	3.23	-0.24	-0.47	4.05	-0.12	1.94	1.13	1.71
EducationStill_Studying	-0.56	0.26	-2.17	-0.55	0.13	-4.09	-0.13	0.14	-0.93	-0.10	0.27	-0.36	-0.19	0.37	-0.51	-0.32	0.30	-1.06
EducationUp_to_15	0.10	0.14	0.75	-0.03	0.12	-0.29	-0.30	0.13	-2.20	0.03	0.26	0.11	-0.10	0.34	-0.29	-0.03	0.20	-0.13
OccupationManual_workers	0.05	0.16	0.30	-0.14	0.12	-1.18	-0.27	0.12	-2.24	-0.60	0.34	-1.75	0.37	0.32	1.14	0.07	0.26	0.28
OccupationNot_working	0.61	0.11	5.78	0.19	0.08	2.35	-0.11	0.09	-1.27	-0.10	0.19	-0.54	-0.03	0.24	-0.11	0.24	0.16	1.51
OccupationRefusal	1.65	0.99	1.67	1.43	0.91	1.57	1.42	0.92	1.53	1.24	1.40	0.89	-0.94	3.40	-0.28	1.00	1.09	0.91
OccupationSelf-employed	-0.04	0.14	-0.29	-0.22	0.10	-2.27	-0.25	0.10	-2.42	-0.17	0.24	-0.71	-0.59	0.37	-1.62	-0.11	0.21	-0.51

6.5 Results for Machine learning

Dataset 1 (WRAP)

The machine learning through the random forests algorithm identified consumer behaviours such as using leftovers, preplanning, discard behaviours and checking cupboards prior to shopping trips as being important variables (Figure S1). In addition, demographic variables such as household size, composition, ownership status and house type were all considered important.

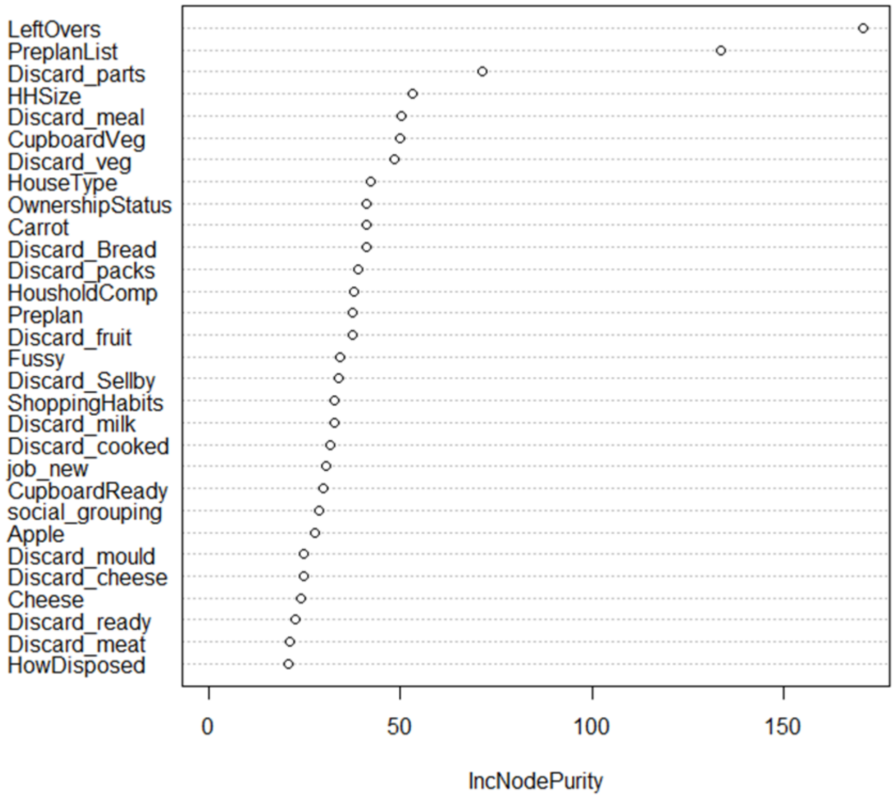


Figure S1. The random forest identified leftovers, preplanning and list writing, discarding inedible parts, homemade meals and vegetables, checking cupboards before buying vegetables as well as house type, household size and ownership status as important variables in classification of the avoidable waste.

The BN identified household size, household composition, how waste is disposed, the presence of fussy eaters, and discarding foodstuff in the last week as important variables (Figure S2).

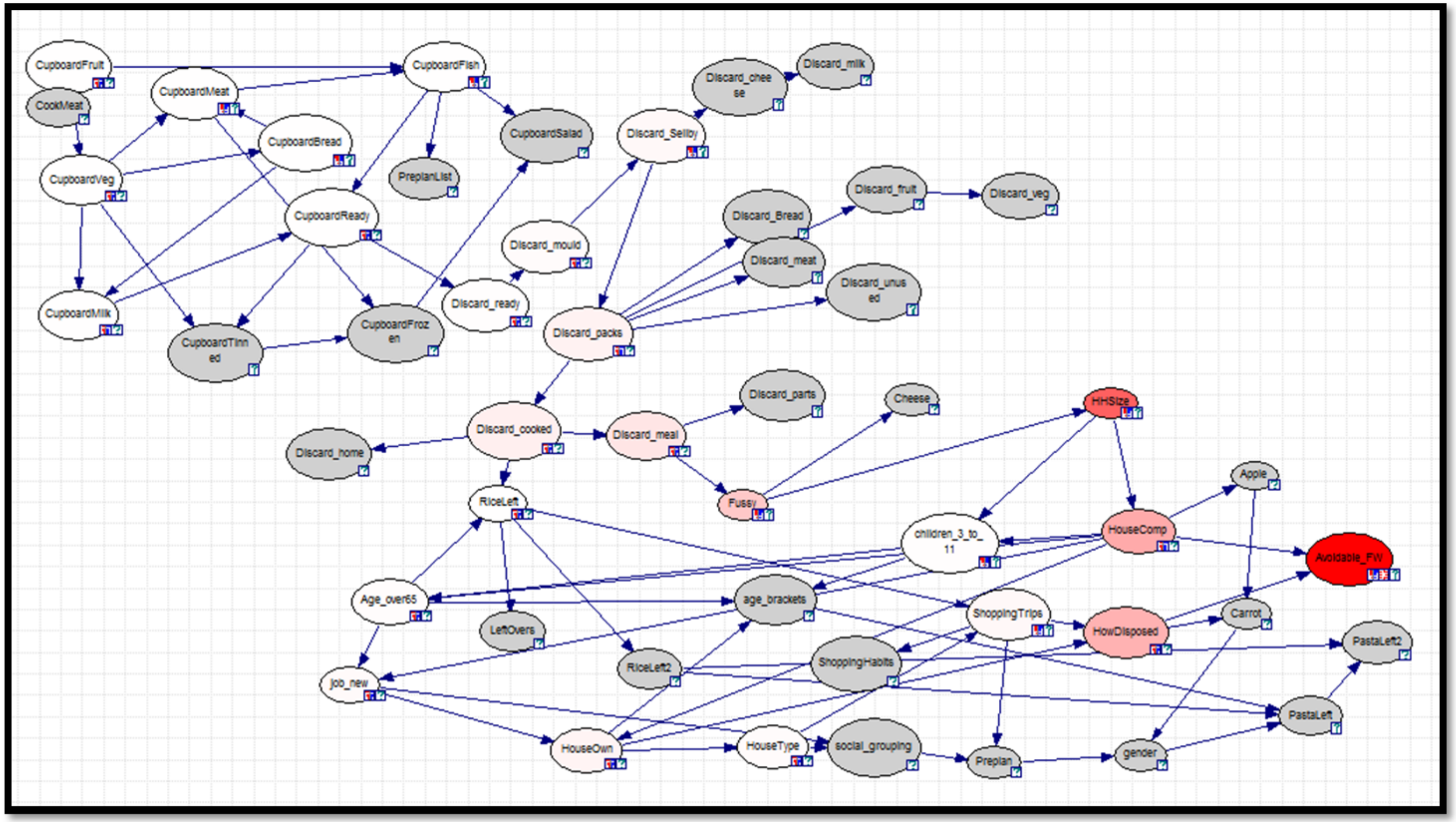


Figure S2. BN structure for Avoidable waste. The node avoidable waste ("Avoidable_FW") was most sensitive to household size, household composition, how waste is disposed, the presence of fussy eaters, and discarding foodstuff in the last week.

Dataset 2 (Flash Euro-barometer 388)

For the Euro-barometer dataset the machine learning through the random forests algorithm identified country as being the most important variables (Figure S3). In addition, age, household size, house management preference, litter in the local environment, agreement with the statement that the family wastes too much and community type were all important (Figure S3).

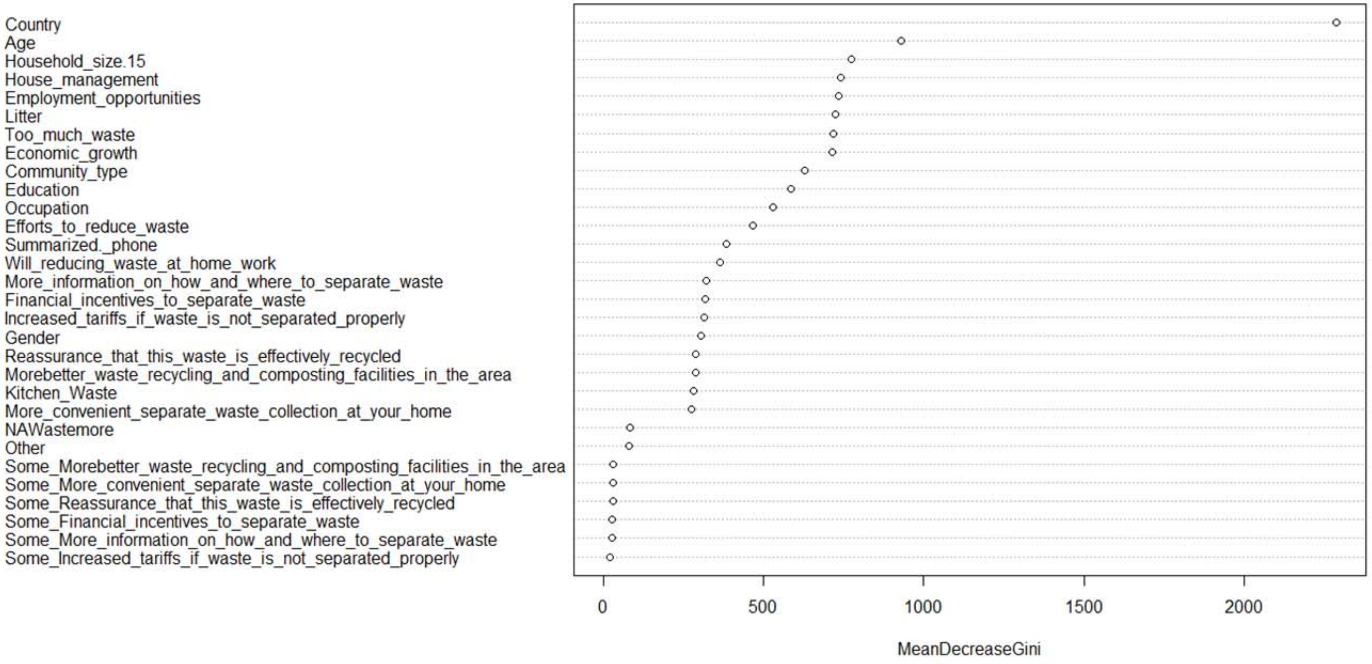
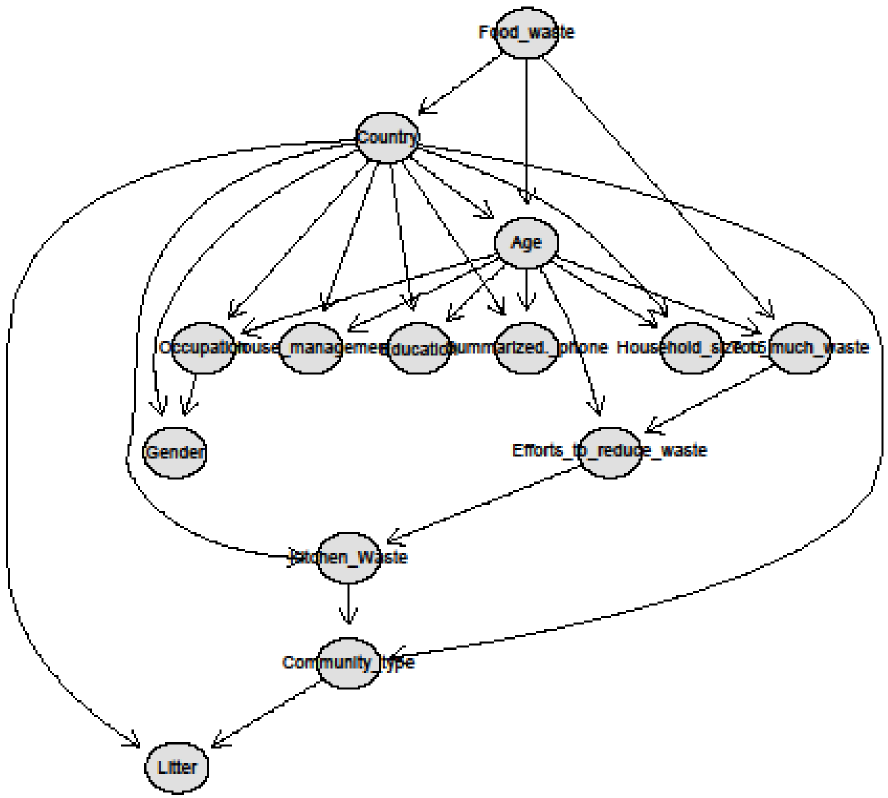


Figure S3. The random forest identified country, age, household size, agreement that employment opportunities for the country stem from waste management, litter and agreement with the statement that the household wastes too much as important.

The Bayesian network showed that changes in the age, country, agreeing that your family wastes too much, education and occupation nodes led to the largest change in the food waste node. The structure of the network is shown in Figure



S4.

Figure S4. BN structure for dataset 2 (Euro-barometer 388) self-reported food waste. The node food waste was most sensitive to age, country, agreeing with the statement “my family wastes too much”, education and employment.

6.6 Additional References

Breiman, L. (2001) Random forests. *Machine Learning*, **45** (1): 5 – 32.

Liaw, A. & Wiener, M. (2002) Classification and regression by randomForest. *R News*,**2/3**: 18 – 22.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Scutari, M. & Denis, J.-B. (2014) *Bayesian Networks with Examples in R*. Chapman & Hall.