D4.2 – Model development and data protocol

Authors

Matthew Grainger, University of Newcastle, UK

Gavin Stewart, University of Newcastle, UK

Simone Piras, University of Bologna, Italy

Simone Righi, University of Bologna, Italy

Marco Setti, University of Bologna, Italy

Matteo Vittuari, University of Bologna, Italy

Document title	D4.2- Data protocol socio-economic modelling
Work Package	WP4
Document Type	Deliverable
Date	30 October 2016
Document Status	Final

Acknowledgments & Disclaimer

Executive summary

The generation of food waste at both suppliers' and consumers' levels stems from a complex set of interacting behaviours. Both Agent Based Models (ABMs) and Bayesian Networks (BNs) are ideal tools for simulation, diagnosis and prediction in the context of food waste generation, because they allow the study of complex systems. This report outlines the methodological framework for each modelling approach.

As empirical data on the drivers of food waste are scarce, there is a need for BNs and ABMs to interact. There are several forms this interaction can take along a continuum from independent modules with limited interaction, to sub-models of a fully integrated model. The most likely form of interaction will see BNs acting as data reduction engines for the ABMs with regard to consumer behaviour and/or for BNs to contextualise ABM outputs.

The contribution of REFRESH Work Package 4 (WP4) to the understanding of food waste generation across Europe depends on the sharing of data among REFRESH partners. To ensure that this occurs in the most efficient way, a protocol provides guidelines for formatting and naming of datasets, defining and describing metadata, and for sharing data.

Table of Contents

List of Figures and Boxes

List of abbreviations

ABM	Agent Based Model
BN	Bayesian Network
DoA	Description of Action
EU	European Union
FW	Food Waste
LCA	Life Cycle Assessment
LCC	Life Cycle Costing
ORP	Other-Regarding Preferences
UK	United Kingdom
WP	Work Package

1   Introduction

WP4 of the REFRESH project implements a behavioural economics approach in order to identify and measure the most important socio-economic conditions and potential policy interventions driving businesses' and consumers' choices in the generation of food waste. More specifically WP4 aims to provide new information on consumer and business behaviour by measuring the effects of major tangible factors of food waste, by identifying hidden and emerging profiles of consumer' and business' behaviours affecting food waste, and by allowing the detection of intangible food waste drivers (Obj. 1). Such an objective is achieved through the development and the testing of Agent-Based Models (ABMs) and Bayesian networks (BNs) (Obj. 2).

The generation of food waste stems from a complex set of interacting behaviours of both food consumers and suppliers. Therefore, a complete approach to the problem requires an analysis of both sources of waste. Both ABMs and BNs provide a modelling approach that fits this purpose, since they allow the study of complex systems. This report provides the methodological background needed in order to achieve the modelling objectives of WP4.

Chapter 3 defines what is a model and the modelling approaches that will be utilized within REFRESH WP4.
Chapters 4 and 5 illustrate the characteristics of ABMs and BNs.
Chapter 6 outlines an ABM aimed at assessing the adoption of innovations for preventing or reducing the food waste generated by companies of the retail sector.
Chapter 7 outlines a BN for characterizing consumers' behaviour with respect to food waste.
Chapter 8 anticipates potential modes of interaction between ABMs and BNs.
Chapter 9 identifies a number of good practices to ensure effective data management and facilitate data exchange.

2   What is a model?

A model is, essentially, a simplified representation of reality. Models aim to capture the essential properties of a system, and to reproduce its entities, relationships and rules. In the context of scientific research, there are many different types of models, ranging from graphical representations of the entities and of the functioning of a system (for example in a model of the "cycle of water"), to fully formalised mathematical models (e.g., a Newtonian model of the revolution of planets).

While a model can represent the reality of a system at an arbitrary level of precision, models need to have theoretical underpinning, and to be as simple as possible in order to provide useful insights, and possibly predictions that can be generalised beyond single cases. To preserve its generality and some capacity to explain, simulate and project the system studied, a model requires several simplifying assumptions.

Within WP4, two main modelling approaches will be utilised. The first is ABM, the second BNs. ABMs will be applied to explore both business-level and consumer-level behaviours, whereas BNs will be predominately used to explore consumer-level behaviours in relation to food waste. Finally, potential interactions of ABMs and BNs will be explored.

3   What is an Agent Based Model?

ABMs are computational systems that simulate at different scales «a number of decision-makers (agents) and institutions, which interact through prescribed rules» (Farmer and Foley 2009). Instead of assuming a (or a small set of) representative agent(s), ABMs describe the individual decision-making process, and the rules regulating the interactions among agents. The model output is the aggregate result (and inherent variation) of the agents' choices and interactions among them over a specified number of time steps, or its final point. Moreover, non-evident phenomena or rules can emerge. This, on the one hand, allows for a large amount of heterogeneity among agents and of the interactions rules to be modelled; on the other hand, it allows the introduction of a well-defined institutional structure. Dropping the constraint of analytical solvability, ABMs can be as complex as required to grasp the understanding of the system studied. Nevertheless, it is important to constrain as much as possible the additional complexity of agents and interactions, to avoid generating models as complex and difficult to understand as the reality they are designed to study.

ABMs allow the researcher to take into account the complex nature of human behaviours and interactions. As such, unlike many standard models, their analytical solution is seldom possible. The typical tool to analyse ABMs is computer simulation. Simulations can take many forms, but the most common is the *Monte-Carlo simulation*, which consists in choosing a given parameter set, and iterating the dynamics of the model many times, with different sequences of random numbers. Such simulation allows the study of the statistical characteristics of the simulation output (means and standard deviations of the results, their distribution, and the occurrence of rare extreme events). By modifying the parameter sets chosen, it is possible, then, to *check the robustness* of the results obtained, as well as to assess the implications of a shift in one of the model parameters. A well-developed model can be used as a virtual laboratory, as it allows the generation of alternative time-series under controlled "quasi-experimental" conditions. As such, ABMs can also be studied with *regression* techniques, exploring the correlations among different parameters and outcomes, as well as the impact of different types of heterogeneity. Given that many relationships among variables are typically hard-wired within the model structure, causation structures can be studied within the framework of ABMs.

An alternative method of analysis frequently used to assess ABMs is the *comparison of scenarios*. Within this method, different initialisations and sets of rules of the model are created to simulate specific known cases (such as two countries), or to study the impact of a policy intervention. Both the aggregate outcomes, and the individual trajectories of development of the agents are then assessed comparatively. The analysis of the results of ABMs frequently relies on *graphical analysis*, such as plots and figures. These are a powerful method of representing both aggregated results and single simulation case studies.

The ABM technique can help generate results and explanations about aggregate socio-economic processes, answer questions concerning the emergence of stable behaviours or behavioural patterns, the stability of a model to small changes, and the evolutionary patterns of the system studied.

3.1 Elements of an ABM

In order to design and develop an ABM, it is necessary to specify at least three main elements (Gilbert 2008): the *entities* (or agents), their *interaction rules*, and the *environment and institutions* within which agents interact.

The *agents* are the autonomous and discrete decision-making units whose behaviour is modelled. In socio-economic simulations, such entities are typically persons, companies, or even nations. Depending on the characteristics of the model, agents can be very simple, or extremely complex, but they usually possess a set of characteristics that can be summarized as follows:

- *Attributes*: the set of agent properties (e.g. age, degree of pro-social behaviour, wealth etc.). Each attribute is typically assigned one (or more) precise value(s) at each given moment. They provide the building blocks to be combined in the rules of behaviour. Attributes can be constants (properties of the individual agent that do not changes during the simulation, such as gender), variable (properties that can change as a consequence of the interaction among agents, or of some external factor), or evolutionary (properties that evolve in time with some endogenous or exogenous rule of change).

- *Rules of behaviour*: they determine how agents behave, and represent the assumptions made about the decision-making processes of the individuals. Agents can, in such a way, maximize some discrete or continuous function using some heuristics, or following some deterministic rules (if x, then y). The rules of behaviour use as inputs agents' attributes, and may also use their memory, and their perception of reality, as well as those of other individuals in the model. The simplest type of agent acts randomly, and is thus named "zero-intelligence agent" (Gode and Shyam 1993). Such type of agent is useful as a benchmark.

- *Memory*: provides the agents with the possibility of recalling past actions and interactions, as well as their results. Memory can constitute a mechanism to condition the behaviour of agents on past actions (for example, cooperate with partners that were cooperative in the past). The simplest type of agent does not have any ability to condition his behaviour on past actions ("zero-memory agent").

- *Perception of the environment*: determines what an agent knows and how he perceives the surrounding environment. In particular, the perception of reality defines the set of information that the individual has when applying its rule of behaviour.

It should be noted that, within an ABM, there are typically many agents, which present different specifications of these characteristics. In this case, the model is called a "heterogeneous agents model". While heterogeneity can provide useful insights into the system modelled, it is important to keep the degrees of heterogeneity to a minimum, thus preserving the explicatory power of the model.

ABMs are characterised by the presence of many interacting agents. Within ABMs, there is usually no central "authority" defining the *interactions*, and these are typically decentralised, with autonomous agents interacting according to their own

attributes, applying the decision rules and using the information available to them.

The *interaction rules* set the constraints on how agents can interact. Depending on the type of model, the interaction rules can be represented in game theoretical form (agents receive a payoff that depends on their actions and on those of other players, according to some table of payoffs), as economic exchanges (one or more individuals buy something that someone else sells in exchange for a good or some amount of currency), or as exchanges of information.

Not all interactions are always possible and, indeed, exchanges typically happen on a defined *interaction space*. The most basic interaction space is represented by a setup where each individual can interact with other individuals without constraints. In socio-economic models, more frequent is, however, the case where interactions are constrained to happen on a specific structure, such as a network or a grid, that can be fixed or evolve as a consequence of individual behaviour. The network structure may represent physical limitations or social/informational constraints. Finally, agents may be mobile and, thus, the interaction space can represent a geographical extension with physical proximity as a constraint.

Beside agents and interactions, the third important element of agent-based models is the design of the *environment and institutions* characterizing the system studied. The environment defines the external constraints and characteristics that influence all agents (or groups of them), and their interactions. In practice, this is the set of all policies, taxes, laws, institutions, and characteristics of the natural environment influencing and shaping the behaviour of all (or groups of) agents. Changing the nature or the incentives provided by the environment and the institutions allows a comparative study of their impact on the agents and on the aggregate outcomes of the model.

3.2 Limitations of ABMs

The problem of Food Waste (FW) generation is a complex one, where agents with heterogeneous behavioural typologies interact in non-trivial ways. Studying this problem through a set of ABMs allows the necessary flexibility to address it. However, there are a few important limitations of this approach that require some considerations. The first is that, with ABMs, it is almost always impossible to formally (analytically) prove the results obtained through simulations. The second is that, as models grow in complexity, it can become difficult to disentangle the impact of the single variables on the aggregated outcomes. When true emergent phenomena (Holland 2000; Wolfram 2002) characterize the model, it may actually become impossible to separate the aggregate results from its individual determinants, as «the whole is more than the sum of the parts» (Anderson 1972). Finally, in many cases, the agents are parameterised using informed estimates, or rough data, and then the accuracy of these estimates is judged by comparing the behaviour of the agents in the modelling environment to the empirically observed one. Such process is quite time-intensive.

4 What are Bayesian networks?

Relationships between variables (both direct and indirect) can be represented by a directed graph. This graphical representation of relationships forms the structure of a Bayesian Network (BN). These relationships are not considered causal (although they can be) but dependant ones. Underlying these dependencies is a set of conditional probabilities (where the state of one variable is conditional on the state of another). BNs are well suited to analyse complex problems characterised by high uncertainty and where a decision is required.

4.1 Elements of a BN

BNs consist of three key elements. These are *nodes*, *arcs* and *conditional probabilities*. *Nodes* represent variables, while direct relationships between variables are represented by an *arc* (i.e. an arrow). Underlying *conditional probabilities* (i.e. probabilistic relationships between nodes) are calculated using Bayesian updating. Nodes can be in a number of categorical states, or can be continuous (although computing mixed distributions of data can be problematic). Conditional probabilities are calculated to determine the probability that node w is in state x given the state y of node z.

Figure 1. Example of a BN from the Norsys Software (1992-2011) Netica.

Top plot: nodes represent variables and arcs (arrows) represent conditional dependencies.
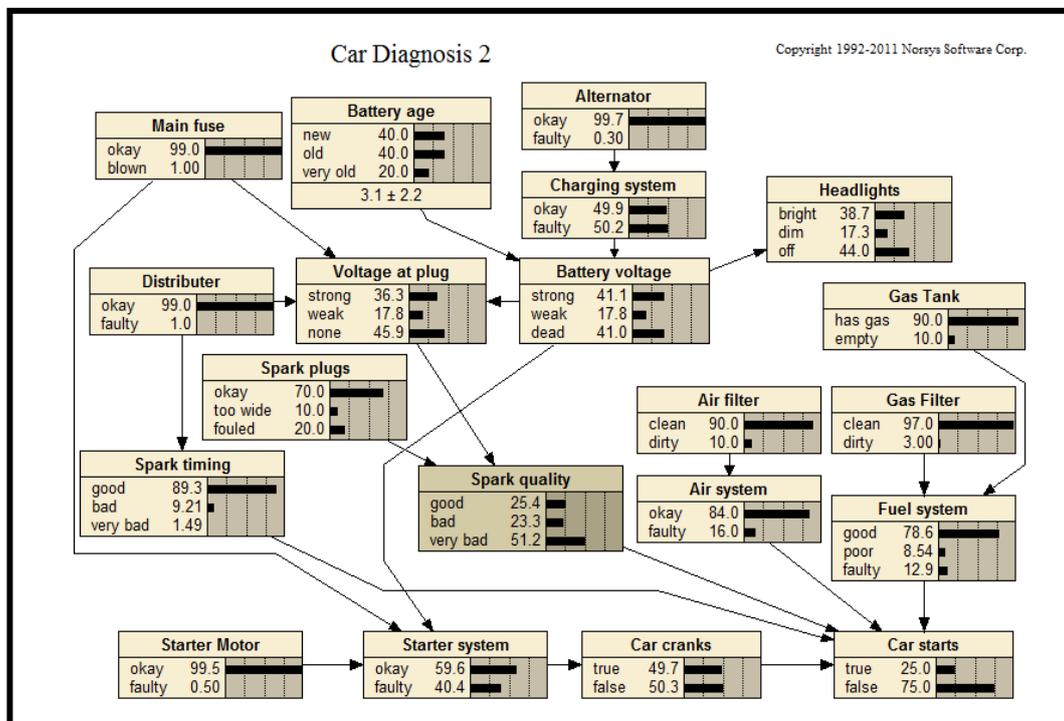
Bottom plots: conditional probability tables underlay the relationships between nodes.
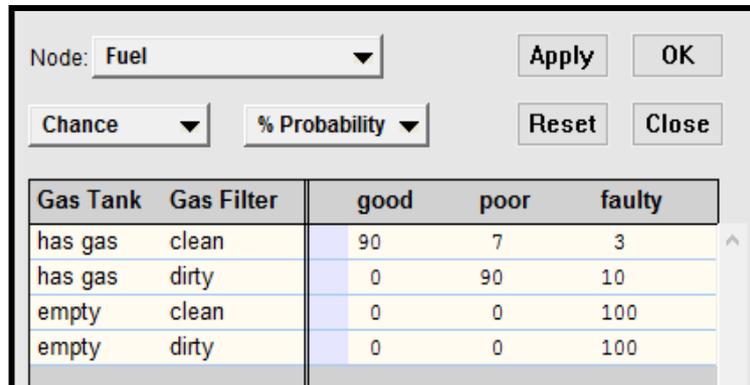
Node: Fuel					Apply	OK
Chance	▼	% Probability ▼			Reset	Close

Gas Tank	Gas Filter	good	poor	faulty
has gas	clean	90	7	3
has gas	dirty	0	90	10
empty	clean	0	0	100
empty	dirty	0	0	100

BNs can be parameterised using empirical data, expert opinions, simulation data, data from literature (effect sizes, known and suspected relationships, etc.), or a combination of these. Machine learning, a subfield of computer science, allows the model structure to be determined through algorithms. These algorithms search for the most likely relationships between nodes, given the data provided. Machine learnt networks can be unsupervised (i.e. the computer searches for the structure without human inputs) or semi-supervised (i.e. a person inputs known or suspected relationships into the model, but leaves the computer to learn the remaining structure). Throughout the REFRESH project, machine learning will be utilised, but pilot work (see D4.1a) suggests that a structure may need to be imposed, thus using a semi-supervised network structure.

4.2 Limitations of BNs

The most important limitation of BNs is that they are acyclic (i.e., it is not possible to have feedback loops in the model). This limitation can, however, be mitigated for by developing linked sub-models, where arcs can cross temporal or spatial scales. For example, a temporal model could be parameterised using several linked models at T, T^{+1}, T^{+2}, $T^{+\cdots}$, etc. The inputs for one model could be outputs from another. This structured object-orientated approach allows the main limitation of BNs to be circumvented.

5 An ABM of food retailers

A model of the food retail market with stochastic networks and community evolution is illustrated within this chapter.

5.1 Economic modelling

Models have been used since the dawn of economic science in order to study the behaviour and the choices of individuals, and their aggregated consequences. Over the years, economic models stratified into a standard economic theory, that largely dominated the discipline and its influence on policy-making. The standard theory of economic choices assumes that individual decision-makers are rational in their preferences, and make their decisions in isolation, considering only their own wellbeing (Mas-Colell *et al.* 1995). Under such assumptions, businesses make efficient investment decisions by considering all relevant economic information, and by weighting all their future consequences correctly (Von Neumann and Morgenstern 1953). Furthermore, within the strict assumptions of the standard theory, economic agents have clearly defined preferences, and give no consideration to the consequences of their decisions on other people's welfare (Mas-Colell *et al.* 1995). Finally, when the behaviour of individuals is scaled up, standard economic theory assumes that decisions are made by a representative agent, whose preferences and decisions subsume all individual heterogeneity (May 1946; Kirman 1992). Such assumption is made in order to obtain models that can be solved analytically.

There are two essential criticisms addressed to standard economic modelling. The *first* comes from the observation that economic decisions are made by human beings, who are not fully rational (Haruvy *et al.* 2007; Armstrong and Huck 2010), whose capacity of maximizing profit is bounded (Simon 1956, and 1957), and whose interrelations are seldom completely anonymous (Jackson 2008). Behavioural economics is a heterogeneous body of literature developed in order to understand, and take into account, the complex factors influencing *real* economic decisions. It investigates the consequences of social, psychological and cognitive factors on economic behaviour. Instead of starting from abstract principles, behavioural economics uses laboratory and field experiments (DellaVigna 2009) and observes the real behaviour of people, that *does not* always concur with standard theory. Adopting this approach, behavioural economics attempts to change the perception of economists about individual preferences and choices. Behavioural economics literature is reviewed – with a focus on the different deviations of businesses from rational, selfish, and isolated behaviour – within REFRESH D4.1b.

The *second* criticism of the standard economic theory derives from the fact that economic agents are diverse, and pursue different scopes with different means. The coordination of different wishes and choices is highly non-trivial. In order to be able to generate analytical models with closed form solutions, scholars assume that individual choices can be considered as the choices of one "representative" standard, utility-maximizing individual, whose choices coincide with the aggregate choices of the heterogeneous individuals (Kirman 1992). As proven by Kirman (1992), the choice of adopting a "representative individual" is unjustified, and can

lead to very wrong conclusions. In the framework of representative agent models, one can reach the conclusion that there should be no trading in financial markets (Rubinstein 1975), no wealth distribution concerns (Hakansson *et al.* 1982), and no role for the government (Milgrom and Stokey 1982). Moreover, the exclusion from the model of local interactions among individuals may cause results to stray away from reality.

In order to overcome these problems, recent years have seen the development of a branch of economics with heterogeneous interacting agents (Gallegati and Kirman 1999; and 2015), whose results have been applied from capital markets to economic policy. The aim of this report is to provide the outline of a model of business behaviour on FW, with particular reference to the adoption of innovations aimed at preventing, reducing or valorising FW. Considering the recent critics to standard economic theory discussed above, the natural method to develop and test such model is through ABMs.

5.2 Relationship with the DoA and model motivation

One of the focuses of REFRESH T4.2 is the definition of a framework for the development of an ABM aimed at assessing the behaviour of businesses with respect to FW, and to the ways of reducing it. One of the most important opportunities to decrease the levels of FW generated by firms is though the introduction of innovations. Therefore, here, the focus is on behavioural factors influencing the adoption of technological innovations by companies.

A relevant share of the FW generated along the food supply chain is produced by food processors and retailers. In order to keep the ABM treatable, it was chosen, however, to separate the two sectors and – at least at a first stage – to study them individually, using two different models. The *first* model studies the retail market, and will be extended to the food processing sector in a *second* model. The present document provides an outline of the model design and characteristics.

The choice of separating food retail from food processing is due to the fact that:

- It allows a clearer understanding of the behavioural patterns and incentive structure of the two sectors, and of the consequences of the interaction among individual incentives;

- It allows to disentangle the effects of single changes to the baseline, which would be much more difficult in a more complex setup.

The choice of starting from the retail market, instead, stems from the fact that this is a simpler sector compared to food processing. Indeed, on the one hand, it is polarized between large and small actors and, on the other hand, it faces directly the consumption phase of the food market, which makes agents (firms) easier to model.

In order to obtain a clear baseline, the model will make strong simplifications, that will be then progressively relaxed. This procedure will allow a clearer definition of the drivers of the results.

5.2.1 Research questions

Following the focus mentioned above, the goal is to create a model to answer the following questions:

- Which are the main conditions and drivers influencing the adoption of innovations aimed at reducing FW in the retail sector? How are the diffusion dynamics influenced by these conditions and drivers?

- Which kind of (clusters/networks) alliances among companies emerge at a meso/macro level?

- How do FW levels in a market change when companies are allowed to create (or to eliminate) ties with each other? How does profit distribution changes as a consequence?

Note that the formation of clusters/networks is observed in reality and is expected in the model as a result of the process of creation and destruction of ties. The latter process is based on companies' material and immaterial payoffs[1].

5.3 Fundamental entities

There are two fundamental entities in this model: companies and (final) consumers.

5.3.1 Demand side

In this model, it is assumed that food consumers (the demand side of the market) are a homogeneous mass of individuals. They are further assumed to be able to absorb any level of offer at a given price (*perfectly inelastic demand*). This modelling choice implies that over-supply and over-demand are excluded from the baseline. More realistic demand functions and segmentations among consumers can be introduced as extensions.

The initial assumptions imply that the quantity purchased by final consumers at any price does not change. However, it does not imply that a company can set any price. Indeed, if one company (in isolation) moves the price above the industrial standard for its size, it loses all clients, since demand is – initially – assumed to be perfectly elastic to price.

5.3.2 Supply side

Sector characteristics

The design of the retailer sector (the supply side of this model) follows from the characteristics of the EU food markets. Here, suppliers operate in a condition of imperfect competition, where large-scale retailers act as price leaders, while small actors are price-followers (European Commission 2014; Perloff 2008; Digal and

---

[1] There is some literature pointing out that retailers are interacting also with processors and producers, not only with peers. This element is assumed out of this model outline, for the moment.

Ahmadi-Esfahan 2002; Sexton 2013). Both types of companies define a full-cost price of exclusion: price coordination and different equilibrium prices result (Sylos Labini model, 1957). The competition is thus based mainly on the quantity supplied. This means that retailers always prefer to supply larger quantities, as this increases their turnover and thus reduces the impact of their fixed costs of production. Finally, in line with the focus of the model, all retail sector operators are assumed to "produce and sell" a homogeneous food product.

Companies are characterized by both their physical characteristics and by different behavioural concerns. Decision making is assumed to be based on both economic and non-economic concerns.

Companies "physical" characteristics

Companies are characterized by five essential variables:

- a given level of unitary costs, to be paid for each unit of good produced: v_i

- a given level of total fixed costs to be paid regardless of the production level: k_i

- a given level of FW (univocally corresponding, each, to a given technology): w_i

- a given level of net production y_i, defined as the difference between the gross production x_i and the level of FW w_i

- a set of relationships F_i^t (the ties shared by a company i with other companies j at time t).

The value assumed by these variables can be influenced by:

- The *size of the company*, which constrains the gross level of production (x_i). In line with the characteristics of the market, company sizes are two: *small* or *big*. Abstracting from considerations about technologies and innovation, the size determines all above-mentioned aspects of the company, except F_i^t (i.e. there are two levels of the above variables: s and b). Further heterogeneity about costs for given sizes could be introduced if empirically grounded, or as an extension to this model[2].

- The *technological level of the company*. As discussed below, a single technology at time (at least initially) is considered. For each company, two types of technologies are available: a basic one, that generates high levels of FW (w_H), and an innovative one, that reduces FW to w_L. The latter can be potentially introduced, but implies a change in the values of the economic variables of the company[3].

Companies "behavioural" characteristics

---

[2] Another important driver of companies' willingness to adopt innovations is their financial situation. Dealing with the financial situation however would mean to add a further layer of complexity and is thus relegated to a model's extension.

[3] See below for a list of possible effects of introducing an innovation, whose characteristics are drawn from the findings of D4.1c.

In line with the findings of literature (see REFRESH D4.1b), businesses are assumed to have three types of concerns included in their utility functions:

- A certain weight (α) is associated to the difference in the selfish economic outcomes obtained as a result of adopting the two possible technology levels.

- A certain weight (β) is associated to the presence of other-regarding preferences (ORP), since companies are concerned with their levels of waste.

- Finally, a certain weight (γ) is associated to imitation, following from the social component of company preferences[4].

For simplicity, it is assumed that the three components are addictively separable, so that the discrete choice is:

$$\max_{w_l, w_h} U = \alpha * \pi_i + \beta * ORP_i + \gamma * Imitation_i(neighbour_i)$$

In the baseline setup, it is assumed that α, β and γ are equal for all agents. This assumption will be then relaxed, allowing both idiosyncratic and group level heterogeneity.

5.4 Timing of the model

The ABM outlined in this document evolves along discrete time periods t. The initial setup includes an unrelated set of N companies, N_B of which are big, and N_s of which are small. At each period t, a series of events happen in sequence. Each company takes the actions of this sequence simultaneously, based on the information available from $t-1$, and on the previous actions occurred at t. The model is run for t_{max} time steps, then stops. The intra-step timing of the model is summarized as follows:

1. New ties among companies are possibly formed;

2. Old ties are stochastically reconsidered for elimination;

3. With probability P_{tech}, the levels of FW (corresponding to the levels of technology) are revised. For revision, companies only observe their own current social networks, and the economic setup at $t-1$ (limited information can be considered as an extension);

4. Levels of production are fixed given the technological decisions;

5. Production (that for retailers corresponds to what is put on the shelves) is absorbed by the demand.

Each of the steps will be now considered more in detail.

---

[4] While it is clear that the bottom line of companies is mainly the creation of profits, bounded rationality in economic behaviour and the ontological uncertainty makes conformity emerging as a possible heuristic, influencing decision making.

5.4.1 Creation of social ties

While considering whether to change technology, companies take into account the current situation (in period t) of their social neighbours (companies that are connected with them). This simulates imitation (in other words, a behavioural shortcut), together with the act of construction of a company's social network.

In the baseline, companies start off as isolates. Then, at each time step t, with given probability P_{link}, each company can create a tie between himself and someone else he is not already connected with. Such ties represent dyadic social connections (whose meaning is detailed in section 6.4.2) between two agents (e.g. between two companies such as Nestle and Unilever, or between Unilever and a local shop). In the baseline, P_{link} is equal for everybody, but it can be made heterogeneous within the extensions. Random creation of ties may seem simplistic but, coupled with the destruction of ties based on their effects, it makes the setup more realistic.

In addition, whenever a group of businesses becomes sufficiently densely connected, they create a coordination-community, that changes their incentives to adopt different technologies, by changing the levels of v or k, and by offering them the possibility of changing market prices (see section "coordination" below). Coordination-communities are, scaffolding structures imposing some constraint or directing the behaviour of groups of agents and having some internal structure (e.g. the European Retailer Organization).

5.4.2 Destruction of ties

As mentioned above, ties are created uniformly at random. Each tie is maintained with probability proportional to its effects (weighted, if there are multiple effects in different directions) on *each* of the interested individuals (companies). Among the possible consequences[5] of tie maintenance are:

- Possibility of modifying the price instead of compressing the profit margin (q') when coordinating with firms covering a sufficiently large share of the market.

- Share the difference in fixed costs (Δk) induced by the introduction of the FW reduction innovation necessary to achieve w_L among the connected individuals.

- As long as they stay connected, less efficient individuals have a reduction of v_i (thanks to the adoption of the learning-by-doing of the other).

- Other possible drivers of link maintenance/destructions (to be identified).

Note that most of elements leading to the destruction of ties are related to economic issues. This happens because the levels of FW are interpreted, here, as by-products of other economic and behavioural choices.

Furthermore, networks could involve also modelled "government bodies" or "associations of category" as "external fields" influencing all partners' behaviour.

---

[5] To keep the model simple, only one combination of consequences is studied at each time.

5.4.3 Choice of creating innovation (of the production technology)

Once the social network has evolved, then each company – with probability P_{tech} extracted independently for each company – decides upon changing its technology level. Companies' investment decisions are strongly influenced by pure economic incentives; however, behavioural economics literature has shown that not only economic incentives matter for company decisions. An extensive analysis of the deviations from egoistic profit maximization is carried out in REFRESH D4.1b. As noted above, companies decide whether to adopt the innovation with a view to maximizing their utility:

$$\max_{w_l, w_h} U = \alpha * \pi_i + \beta * ORP_i + \gamma * Imitation_i(neighbour_i)$$

where α characterize the degree to which a company is driven by profit motives, β the importance attributed by the company to the level of waste (or, in other words, the relative importance of other regarding preferences within a company's utility) and γ the importance attributed to the behaviour of other individuals surrounding the company.

While, for simplicity, companies are initially assumed homogeneous for what concerns their behavioural characteristics, heterogeneity will be introduced as an extension. Both the specific types of ORP (pro-environmental behaviour) and social interaction characteristics (conformity bias) to be implemented, and the relative levels of α, β and γ are drawn from the literature review, using hints from other REFRESH WPs. In particular, α, β and γ will be set as "country representative", in order to compare the situation in different countries. National case studies should provide evidence of their relative strength in different countries.

The utility function of companies will be operationalized as follows. Company choice is a discrete one, between the continuation of producing a high level of waste ($w_i = w_H$), and changing the production technology, thus reducing the level of waste ($w_i = w_L$). The company chooses the option which gives the maximum utility between:

$$U_H^i = \alpha(p_H * x_H - v_H * x_H - k_H) - \beta\left(\frac{w_H}{x_H}\right) + \gamma([\#w_H \in F_i]_{t-1} - [\#w_L \in F_i]_{t-1})$$

$$U_L^i = \alpha(p_L * x_L - v_L * x_L - k_L) - \beta\left(\frac{w_L}{x_L}\right) - \gamma([\#w_H \in F_i]_{t-1} - [\#w_L \in F_i]_{t-1})$$

waste level w_H is chosen if $U_H^i \geq U_L^i$ while w_L is chosen if $U_H^i < U_L^i$. The direction in which all other economic variables move depend on the type of innovation studied (see subsection on the impact of innovation below).

The equation above assumes that firms are somewhat pro-environmental (hence, food waste reduces their utility with proportion β) and tend to conform to their neighbours' choices (third element of the utility). Other types of utility functions can be explored, with different combinations of ORP and other social mechanisms.

It should be noted that, with the setup presented, in principle, a technology can be adopted and dropped repeatedly if the situation is borderline. This unrealistic

outcome is avoided by assuming that – for a given company – technological decisions do not happen at every step, but with given probability P_{tech}.

5.4.4 The impact of innovation

Company decisions regarding which technology[6] of production to choose and whether to adopt or abandon an innovation depend, in large part, on the financial return of the investment in that technological innovation. For this reason, the innovations will be considered in a highly-stylized form, being only represented by their impact on FW. For simplicity, it is assumed that only one possible innovation is available at a time (competition among technologies will be addressed as an extension to this work).

Given their structural typologies (small or big), businesses can produce two levels of waste: low (w_L), or high (w_H). The level of waste produced depends on the technology adopted. In practical terms, the technological choice may have an impact on v, k and x in different ways:

- Diverging effects on k and v, while x is increased (decreased). The innovations allow FW set to w_L (it follows that, with this innovation, y is increased).

- k is decreased, but FW becomes w_H (environmentally detrimental innovation).

- k is increased steeply, no effect on v, y is increased, while FW becomes lower (w_L)[7]

- v is increased steeply, no effects on k, y is increased, while FW becomes lower (w_L)[8]

- Other types of innovation, as well as examples of innovations, following this setup, will be drawn from case studies, focus groups, and from the literature review of REFRESH D4.1c.

5.4.5 Coordination

If their total production is above a given threshold of market quota (λ), companies that are in the same *connected network* (i.e., that are directly or indirectly connected to each other) can modify the price (instead of the margin of profit q'') when passing from one technology to another (more expensive in terms of costs of production). Given the network-creation mechanism, this can happen only if the companies gain positive benefits from sharing a link, or if some degree of trust develop among them (so that they effectively coordinate on raising the

---

[6] Technology is here intended in its economic meaning, thus comprising both proper technological innovation and organizational innovation. Indeed, innovations are only considered for their impact on FW and on economic variables.

[7] This specific type of choice is not particularly relevant for retailers, but could be useful when dealing with producers.

[8] This specific type of choice is not particularly relevant for retailers, but could be useful when dealing with producers.

price, and do not free-ride on each other[9]). Furthermore, differential levels of co-ordination can be included for directly and indirectly connected individuals (as an extension).

In presence of higher levels of network density, the network can have a phase transition and become a *coordination cluster* (or a *community action*). A cluster can coordinate and move together from one technology to another systematically. In order to provide them with a device of coordination, clusters are assigned a voting system (the comparative effect of different voting systems can be assessed in extensions), employing a simple majority rule to decide upon technology.

5.4.6   Choice of unitary prices and supply level

Given the characteristics of the food retail market, the price is fixed, with a mark-up on the full costs of production in line with Sylos Labini (1957) model. Thus, each company fixes a margin of profit (or mark-up) q'', and unitary prices to:

$$p_i = v_i + \frac{k_i}{x_i} v_i + q'' v_i, \text{ or equivalently: } p_i = v_i + \frac{k_i}{y_i + w_i} v_i + q'' v_i$$

Given the above discussion, the only decision variable here is x_i. But since, in the food market, companies compete prevalently on quantities, x_i will assume maximal one for given technological level and company size. It can be concluded that, unless coordination among firms happens, p_i is completely constrained.

5.5 Observables

The model starts with no company having adopted the technology that allows w_L, and studies, given the parameter set, the patterns of evolution of innovations adoption, as well as its asymptotic values.

Within this model setup, it is possible to derive testable simulations relative to:

Observables	Validated (through expert opinions)
Reduction in the level of food waste (by categories of companies, sector, and overall).	Proportion of food production/consumption that no longer results in waste at the level of single company/consumer.
Degree of diffusion of a given technological innovation.	For an already adopted innovation, the proportion of companies in the area that uses it; For a brand new innovation the poten-

---

[9] Explicit price coordination is illegal in the EU, however a raise in price, introduced by competitors to compensate economic cost of innovations can be an indirect result of coordination actions with other aims (e.g. fostering the diffusion of an innovation) and can result in positive outcomes from the FW point of view.

	tial (declared) interest in deploying it.
Threshold of innovations impact to allow the emergence of imitation//cooperation/collusion among companies;	Qualitative assessment tracking phase transitions in the diffusion of an innovation in pilot countries.
The (distribution of the) proportion of a connected group needed for it to engage in innovation introduction.	Qualitative assessment studying the coordination actions and courtyard agreements.

These observables are to be studied comparatively, as function of:

- The economic setup of the industry;
- Behavioural choices (other regarding preferences and behavioural heuristics) and their relative importance;
- Types of innovations;
- Types of policies implemented.

5.6 Policies

Many different policies could be implemented in order to try to reduce FW in the described setup. In the context of this model, policies are implemented simply as exogenous changes influencing all agents. The first classes of policies studied will be:

- The exogenous creation of clusters (created by policy, with moral suasion), from the beginning of the simulations;
- Exogenous changes in the cost structure of the companies (or of certain categories of companies);
- Exogenous changes increasing certain effects of the innovations (on every agent or only on certain categories).

5.7 Outlook of expected results

While the ABM to be developed within WP4 is expected to produce emerging properties that are difficult to predict, it is possible to outline some of its expected results. Indeed, a (significant) FW reduction innovation is expected to be adopted even if it implies a competitive disadvantage (market contraction) by:

- Answering to (some) consumer demands (when no ties are available);
- Coordinating the decision with other companies through networks and public incentives to network creation (that increase the probability of network stability). This is the case if the role of network ties among firms (and the consequent imitation) prevail over individual other-regarding preferences (e.g. al-

truism), and when the threshold for a coordinated network will be reached at a lower level than the individual threshold;

- For the small companies, acting as followers of big ones, by imitation.

From the micro-economic perspective, an innovation adopted by a retailer to reduce waste, also reduces the demand, the supply x_i (it can increase the fixed or variable costs) and the profits.

Indeed, in line with Sylos Labini (1957), individually increasing the price above the market level (set by "large" companies) can encourage the local entrance of new competitors, and induces consumers to shift to other producers. Accordingly, this means that profits are reduced by innovations addressing FW; hence, supply side economic incentives alone push the companies not to invest in FW reduction innovations. However, if prices could be increased, then the investment could become feasible or even advantageous. In this sense, the coordination (collusion) devices created by the *communities* discussed above can create modalities through which a set of companies act in concert and, thus, invest and together increase the prices.

5.8 Extensions

As frequently noted, this model makes a strong simplifying assumption in order to create a simple baseline against which to compare extensions in multiple directions.

Drawing from the model description above, possible extensions could be:

- Introducing a more realistic demand function, with category of consumers with different preferences for "green" products, and different perceptions of company reputation (assuming that companies can increase their reputation by shifting to the low FW technology).

- Introducing different coordination mechanisms among individuals directly or indirectly connected, and experimenting different voting mechanisms within the coordination-clusters.

- Introducing competition among different innovations introduced in the market at the same time, or in different moments.

- Introducing heterogeneity of the individual behavioural characteristics of companies (α, β, γ).

- Introducing heterogeneity in the probability of network ties formation P_{link}, or directly introducing different mechanisms of network formation altogether.

- Relaxing the strong correspondence between size, production level and ex-ante costs.

- Limiting the possibility of acquiring information about the adoption of a technology by other companies, by limiting it to directly-connected companies.

5.9 Inputs from other WPs and other activities

Task 2.1 is producing an inventory of alliances for FW reduction and valorisation, that could be of interest for WP4. Indeed, it is a stated purpose of this task to elicit three aspects of existing alliances: incentives involved, success factors, and barriers to action. Of particular interest for the ABM are quantitative statistics which will be gathered during the interviews with the leaders of alliances on these aspects:

- Periods of activity (i.e. their length of activity),

- Objectives of the alliances,

- Types of companies involved (dimensions, efficiency, levels of waste production, unitary production costs, fixed and variable costs, styles of management, reputation among consumers, levels of commitment on specific issues);

- Genesis of the alliance and role of the government in its establishment;

- Sectors of activity (position of the alliance in the food supply chain),

- Factors contributing to the success of the alliance (expert opinion), including data on failed attempts to create alliances and on alliances that no longer exist.

Due to the stakeholders engaged, the pilot studies of *WP2* will cover, due to the stakeholders engaged, a very significant share of the retail sector, corresponding to an average of 80 percent of the national markets in the four countries, as well as a sizeable but more limited part of the food manufacturing sector. It therefore makes sense that WP4 starts modelling the retail sector by taking into account the findings of the pilot studies.

The data collected within *WP2* will be of qualitative type. Therefore, they will provide suggestions on the design of the ABM, and on the general direction of the main actions for FW reduction and valorisation observed. It is very important that some comparable characteristics of the players involved are collected and shared. Making an assessment of the business model is otherwise impossible (there are no comparable differences on which to evaluate the outcomes of the model).

WP3 will provide a set of policy options, some which will be tested within the ABM. If WP3 would be able to characterize the impact of policies on categories of players of both the supply and the demand side of the market, it would be of great interest for the design of the scenario analysis within WP4. Connecting the expected impacts of policies with national characteristics would also help the scenario analysis of WP4.

WP5 will collect systematic quantitative data on the effects of FW Management and Valorisation (FWM&V) choices on firms' economic outcome. This data will provide a useful input to for the ABM. While WP4 works on a more macro level than WP5, these data will be useful in defining the impact of innovations on economic actors. In particular, this data will be useful in suggesting the agent characteristics influencing a given FWM&V approach, as well as the direction of influ-

ence. As the most important information for the ABM will be collected in the framework of Task 5.3, where the LCC and LCA are actually performed.

The main interest of WP4 is to identify the systemic economic impacts of innovations and policies, while the collaboration with *WP5* offers the possibility to include environmental impacts in WP4 models. However, environmental impacts warrant a deeper study being carried out later in the project. Furthermore, data from *WP5* could be used in model validation, to assess whether the stylized companies designed in the ABM provide the same reaction as expected given the results of the LCC analysis.

In the future, a possible development of the WP4 supply side model may include some aspects of the internal company structure. Within such an advanced model, data on the LCC will provide WP4 the necessary information to design relevant agents internal characteristics and interrelations. Furthermore, the same information would be pivotal in validating the modelling choices concerning company internal structures.

Task 5.4 envisions the creation of a "Higher system level model", whose declared objective is to analyse the environmental impacts and LCC of possible socioeconomic policy and technology changes, via integrating relevant results from all WPs into relevant scenarios. This model constitutes a parallel development with respect to the model developed in Task 4.5. In particular, the ABM focus will be on the impact of interactions among economic subjects, and on the relationships between them and policies or innovations using hypotheses on agents' behavioural biases. A comparison of the results obtained by the two models will be mutually reinforcing and provide further result validation.

Task 6.1 has identified a list of priority waste streams for the EU-28 based on product sales volume and environmental impact. The identification of such products will help select the market on which the ABM will be applied and focused. Furthermore, *Task 6.3* could provide WP4 with an assessment of the economic impact for businesses of present and future technologies and approached for waste valorisation. Data should include information on company size, turnover, and current waste levels.

5.10 Outputs to other WPs and WP4 activities

One of the main objectives of *WP2* is to help organizations identify the win-wins that can only be achieved by collaboration across the system, for example the different contributions to whole chain resource efficiency. WP4 could provide scenario analyses, making it easier for stakeholder to assess the fruitful effects of cooperation. Furthermore, WP4 could provide inputs on conditions to actually reduce and valorise FW in the presence of different sets of agents or market characteristics. In other terms, WP4 could help compare (qualitatively) the chances of success of different strategies in different situations. Methodologically the construction of the Decision Support System in *Task 2.5* would benefit from the inputs of WP4.

As for *WP3*, the ABM developed within WP4 will help assess and characterize the systemic consequences of different policies. Indeed, while it is possible to control

the direct impact of a policy on a subject, the policy could also have indirect effects due to the interaction of this subject with other stakeholders. Furthermore, with information about the intended or observed impact on single players, the ABM will help assess comparatively different policies, or the same policy in different countries (or in areas with different market structures and/or socio-economic characteristics). This will allow to provide both an ex-ante assessment of a proposed policy and an ex-post dissection of the individual and interactional causes of its success (or failure). This, in turn, will help calibrate in more detail the policy recommendations, conditioning them on the presence of certain social and economic characteristics of the consumers or of the supply side market structure.

The main contribution of WP4 for *WP5* consists of providing market-level indications on the implementation of given FW Management and Valorisation (FWM&V) approaches. Through the ABM, WP4 can assess how the systemic impact of a given FWMV is affected by exogenous changes in the market conditions. This would provide a robustness check of action effectiveness under different economic scenarios.

Finally, the ABM of WP4 will contribute to elaborate a systemic perspective on the assessment of valorisation approaches foreseen within *WP6*, stressing the consequences at the industry level of different strategies of waste valorisation. The ABM will further provide an analysis of the large-scale impact of technologies diffusion and the approaches to it, and suggest possible paths for encouraging different valorisation approaches.

5.11 Adapting the model to study the processing sector

Adapting the setup described to the analysis of the processing sector require essentially one major change to the setup described. Indeed, it is necessary to introduce a production through which inputs are transformed into outputs with given efficiency. Innovations in this sector can also impact the efficiency of production, thus benefiting the processor itself. The introduction of such changes softens the tension between the economic objectives of the company and the decrease of FW levels. Furthermore, types of innovations that are not relevant for retailers can be usefully studied in the processing sector, for example:

- k is increased steeply, no effect on v, y is increased, while FW becomes lower (w_L);

- v is increased steeply, no effects on k, y is increased, while FW becomes lower (w_L).

6   A BN model of consumer food waste

6.1   Relationships with the DoA and model motivation

One of the aims of REFRESH T4.2 is to develop a framework for the elaboration of a BN to assess consumers' behaviour with respect to food waste, and potential interventions to reduce it.

The waste produced at the consumption stage is thought to be responsible for the largest proportion of food waste in developed countries (Parfitt *et al.* 2010). For example, Stenmarck *et al.* (2016) estimate food waste in the 28 EU countries (extrapolated from data for 11 countries) at 87.6 ± 13.7 (95% CI) million tonnes per year, with 46.5 ± 4.4 (95% CI) million tonnes coming from households. This equates to between 46.7% and 63.5% of the total food wasted. Hence, address-ing this section of the food supply chain will potentially have a proportionally large effect.

BNs produced during T4.1.2 have identified the key drivers of consumers' food waste in a geographically limited area using an objective measure (WRAP waste compositional dataset for the UK) and in a geographically wide area using a sub-jective measure (EU Flash Barometer 388 dataset with self-reported waste). These two BNs confirm the high levels of variability in food waste, and the complexity in the variables driving the levels of FW. Self-reported levels of FW across the EU vary across countries and with age and education status of the respondent. In the UK, the weight of food waste per household varies with household size, as well as the age of a respondent (interacting with household size).

These findings validate the approach of WP1 to assess food waste across several EU countries, as differences among countries are an important factor in driving the levels of reported food waste. In addition, these findings provide important typologies that can form the focus of the questions/discussion within WP1 and help parameterise further models (ABM, BNs, and others) that WP4 will develop as the project unfolds.

In contrast to the ABM model outlined above, the data to parameterise a BN model on consumers' FW are yet to be collected and hence defined. These data will be collected by WP1 and when they are available, several options of modelling approach can be explored. These options are dependent on several factors, for example, if country level differences are, as expected, significant drivers of food waste, the development of country-specific models will be considered in the first instance. These country-level models will be integrated as object-orientated mod-els, so that a single overarching BN can be developed as a decision model for the EU. This overarching model can, then, account for country-level differences, and allow country-level (or country groupings) to be the target of policy interventions (in the model).

6.1.1 Model aims

The BN(s) will aim to:

- Identify the main drivers influencing consumer food waste;

- Identify country-level (or focus group level) differences between drivers;

- Evaluate the importance of the model structure to the understanding of the drivers of consumer food waste;

- Integrate decision relevant nodes into the model, to allow decision support across the EU.

6.2 Fundamental Entities

BNs require three fundamental entities: a measure of food waste (this can be objective, e.g. weight of food waste, counts of food waste components, etc., or subjective, e.g. self-reported food waste), a number of covariates (demographic and behavioural variables), along with a structure (with underlying conditional probability+ structure). The structure can come from: (1) Machine learning (semi-structured or fully automated learning), and/or (2) theoretical constructs of consumer behaviour.

6.2.1 Food waste measure

In REFRESH D4.1a, the inconsistency of self-reported food waste was highlighted. Self-reported data can be biased (Priefer et al., 2013) and therefore our preferred measure of food waste would be objective. However, given the practicalities of sampling many EU consumers, and the costs associated with obtaining objective measures of food waste, self-reported measures of waste have to be used in most cases. WP1 will be calibrating self-reported waste with objective measures in small scale experiments. These experiments will allow us to adjust or to convert self-reported measures of waste to approximant waste levels, if needed.

6.2.2 Demographic and behavioural covariates

A number of behavioural and demographic covariates are required in the BN. These must include variables that have been highlighted as important drivers of food waste (e.g. household size, age structure, employment, planning and shopping behaviours, etc.).

6.2.3 Model structure

Model structure will be determined using machine learning and based on WP1s factors: motivation (attitude, social norm, awareness), ability (knowledge, skills), and opportunity (time/schedule, material, infrastructure), as well as demographic variables. The Bayesian Information Criterion (BIC) will be used in order to assess whether the structure is appropriate.

6.2.4 Additional Entities

Decision nodes can be included into the BNs to identify which interventions are most effective (and/or cost-effective).

Data from WP3 and from the ABMs developed through WP4 will be used to develop decision nodes. Information on management costs associated with each intervention will be useful to allow the full parameterisation of the BN and the development of a true decision model that can inform decision support tools. Information will be shared between the ABM(s) and BN(s) (see Section 8 on the interaction of models) to allow further model development and refinement.

6.3 Timing of the model

At their earliest stages of development, BNs will not have a temporal component. If modelling (both ABM and BN) and assessment indicate the need for a temporal component, then this can be developed using object-orientated models. Temporal BNs may be required in the context of the development of policy relevant decision tools. Including temporal trends in household size for example (increased probability of single person households) may be useful to allow predictive model development, and to test scenarios of change.

6.4 Outlook of expected results

- Key drivers of consumer food waste identified at country and EU levels

- Structure of the conditional dependencies underlying the consumer food waste system identified and tested empirically through BIC interventions identified and linked in to the BNs across scale of the EU (Note: interventions may differ in effect at different country or other scales and this can be modelled in the BN).

- A functional decision support tool allowing scenario testing across the EU (investigating projected changes in demographics or policy environment where appropriate)

- An underpinning probabilistic framework to a road map for European Food waste (T4.5.2)

6.5 Inputs from other WPs and other activities

WP1 will provide the main data (from focus groups, and small scale experiments) necessary to parameterise the BNs. In particular, Task 1.3 involves collecting data to compare different food waste measures. This data may allow to understand some of the biases present in the self-reporting measures commonly used, and to estimate correction factors.

WP3 will provide a set of policy options, some of which may be relevant to consumer behaviour. These can be integrated into the BNs as decision nodes, and the impact of these policies on decreasing food waste can be simulated. Scenarios of future changes in policies or incentives to reduce waste can also be run in the BN.

6.6 Outputs to other WPs and WP4 activities

WP2 aims to identify the "win-wins" that can only be achieved by collaboration across the system. BNs, as probabilistic systems models providing transparent, robust and repeatable decision analysis to determine what node states lead to the optimum solution to allow decreased food waste. BNs allow an oversight of the whole system and its interactions, rather than only the areas stakeholders are familiar with. This helps develop systems-thinking and solutions to systems problems, reducing the potential for unintended consequences of decisions made in isolation.

6.7 The building blocks for a road map to reduce European food waste - *Preliminary ideas*

BNs developed throughout the REFRESH project will address aspects of consumer behaviour in relation to food waste. These networks will eventually integrate potential policy interventions; therefore, they can be used in decision analysis. In combination with ABMs, BNs form the analytical basis for the road map for reducing EU food waste. The road map will include a Graphical User Interface (GUI) that will allow data to be linked spatially to a map of the EU. As shown in Figure 2 below the interface can be used to identify potential areas of high food waste behaviours based on the BN outputs.

Figure 2. The Road map for European food waste GUI.
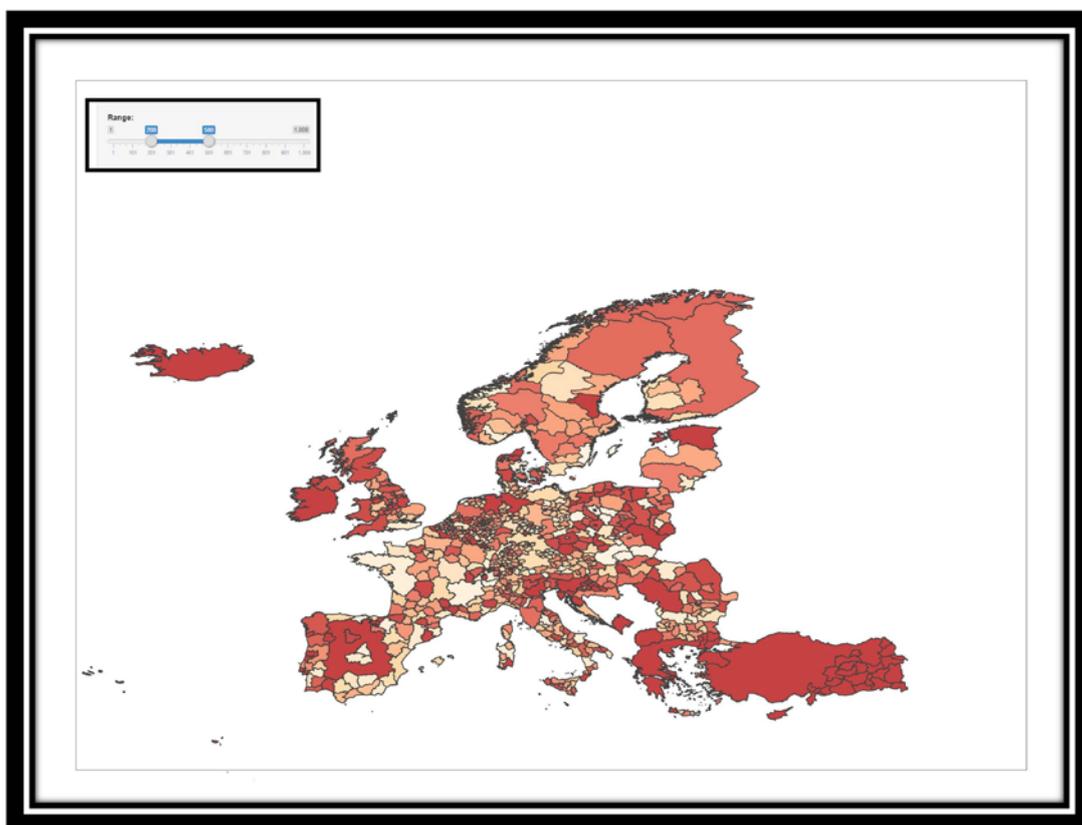
The GUI can be used to predict the probability of high consumer food waste based on demographic variables. Using the concept of conditional probability behavioural responses of consumers and predicted food waste can be mapped on an EU scale. The map in Figure 2 shows a prediction of high food waste based on the household size in 1884 European NUTS (Nomenclature of territorial units for statistics) regions. Regions shaded darker red are more probable to have high levels of food waste. These regions could be important targets for food waste policies.

7   Model interaction (ABMs, BNs)

Given the lack of empirical data and the difficulty of building datasets on food waste behaviour of companies and consumers there is a need for BNs and ABMs to interact dynamically. BNs can be efficiently used as a tool to deepen the understanding of the results of ABMs. Indeed, the ABM model on business FW behaviour requires setting a certain number of behavioural, structural and economic variables, each combination of parameters is subject to yield a certain set of outcomes in terms of adoption of innovation, levels of waste and economic outcomes. The dependence of the latter from the input variables can be difficult to assess especially with reference to the micro-mechanisms involved in the model. BNs can be used to construct probabilistic maps allowing the internal micro-mechanisms to be uncovered, linking the outcomes with the inputs. Given the importance of interaction here the potential forms that it can take are discussed.

In the first form, ABMs and BNs will form independent modules, with inputs to one of the models stemming from outputs of the other, and potentially *vice versa*. For example, BNs are effective at data-reduction and, as such, a BN could provide inputs for the ABM (Pope and Gimmlet 2015). Outputs from the ABM could be used to further refine inputs to the BN (Figure 3).

Figure 3. First type of interaction between ABMs and BNs.
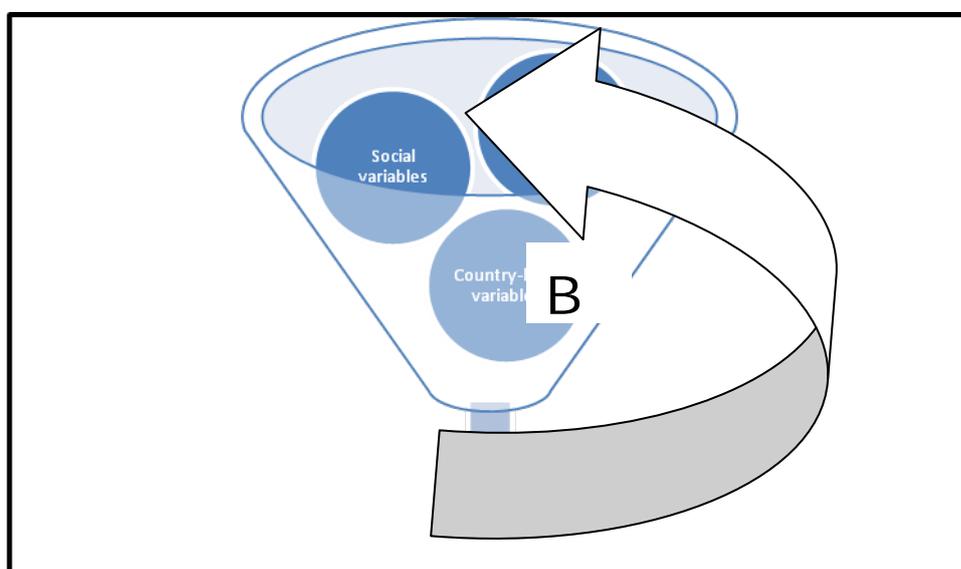
The BN can identify the most important variables and these can become the input for the ABM. Potentially outputs from the ABM could then influence the BN in a number of loops, until model stability is achieved.

In the second form of interaction, the two models can form separate modules of a larger model (see Kocabas & Dragicevic 2013, for an example). Outputs from ABMs and BNs could be used to further parameterise another model (an ABM, a BN, or other model type; Figure 4).

Figure 4. Second type of interaction between ABMs and BNs.
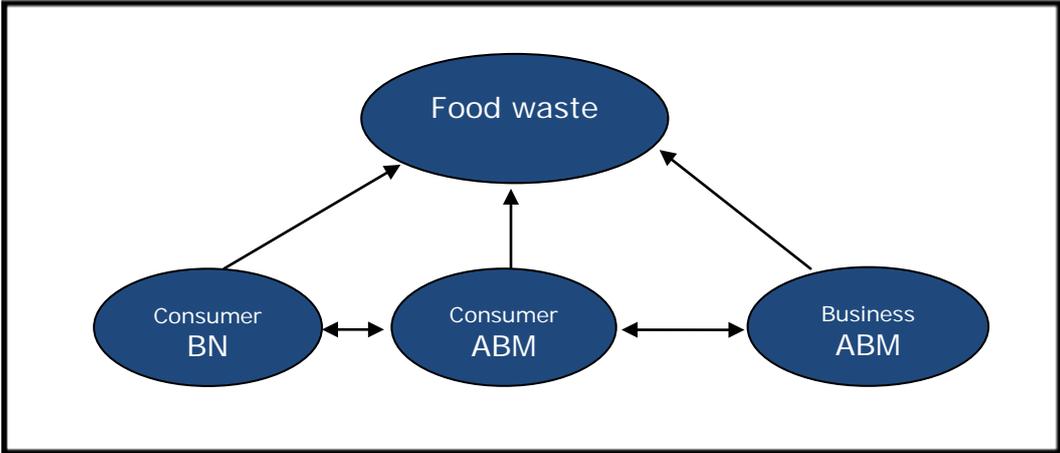
BNs and ABMs can act as independent models feeding outputs into a main master model that addresses a larger scale issue. BNs and ABMs can provide inputs and receive outputs from each other.

In the final form of interaction, the two models can be fully integrated (Figure 5).

Figure 5. Third type of interaction between ABMs and BNs.
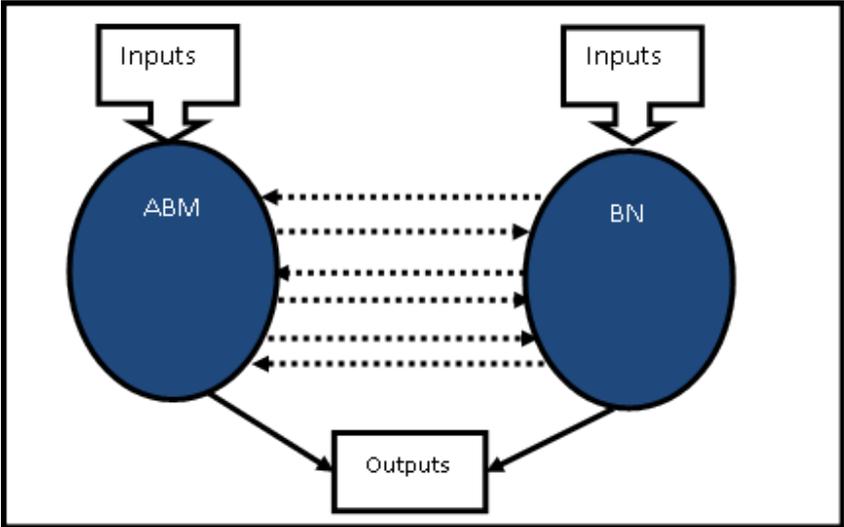
BNs and ABM can have separate inputs but then interact at stages within the modelling process, providing inputs and receiving outputs from the other model. Separate but related outputs can be pooled once model stability is achieved.

The approach adopted depends on the behaviour of the modular components. Where influential interactions occur, integration is required, or the dynamic nature of the dependency needs to be captured in a further model. Important interactions between model components will be specifically examined, and when each modular component is finalised, the integration needs for decision support and the integration methods will be chosen. These decisions will be revised annually and discussed with *WP3*, in order to ensure the utility of the final decision support tool/roadmap.

8   Data standardisation protocol

8.1 The need for a protocol

The sharing of data within and between Work Packages is imperative to the success of the REFRESH project, as recognised by the European Commission's Directorate-General for Research and Innovation (DGRI) (2016)[10]. Effective data management will reduce the costs associated with manipulating data formats, reduce the potential for misunderstanding data (particularly where data are analysed by a different set of researchers from those that collected it originally), and allow future data synthesis and further analysis well past the lifetime of the REFRESH project itself.

The EC DGRI (2016) highlights the need for a data management plan which addresses, for each dataset: naming of datasets, description of the datasets, description of metadata and standards, description of how the data will be shared (inclusive of embargos, software, repository, etc.), and description of how the data will be archived. Herein, this structure will be expanded with specific reference to the datasets for use within WP4.

8.2 Format and naming of datasets

In order to promote data sharing, a format is required that will allow use in a variety of statistical packages and computer programs. In addition, due to the rate at which computer formats become obsolete and/or degrade, the dataset must also be resilient. Data saved as ".txt" or ".csv" are the most universal (in terms of the number of programs/packages that can read or convert them) and the most resilient to degradation.

In some programs/packages, datasets with long (greater than 30 characters) names or with spaces in the dataset name cannot be read. Dataset names need to use only English characters (i.e., no special characters, like @, ~, *, etc.) and have a date in the name, so that researchers can ensure that they are assessing the latest iteration.

8.3 Data definitions and description (metadata)

When researchers collect data, they often use abbreviation or shortened names to identify variables. These can be difficult to interpret if one was not part of the team which implemented original data collection. To aid data interpretation, all variables used need to be defined in full. The dataset needs to be accompanied by a separate file entitled "readme.txt", which gives a brief description of the data collection (including the original purpose of the data collection, the methodology used to collect them, links to external resources that may help the user understand the data, etc.). The dataset should also include any known issues with the

---

[10]   Guidelines   on   Data   Management   in   Horizon   2020 (http://ec.europa.eu/research/participants/data/__ref/h2020/grants_manual/hi/oa_pilot/h2020-hi-oa-data-mgt_en.pdf).

accuracy or precision of the dataset, any information of restrictions to sharing the dataset and any required citation (Box 2).

Underneath this, all variables used in the dataset need to be fully defined (i.e. the metadata). A definition should include: the name used in the dataset, a short description of the variable, the units the variable is expressed in (including all codes used, if appropriate), and any other information that could help the user understand what the variable refers to (e.g., links or citations of sources, techniques or tools used to gather data, temporal or spatial resolution, precision, accuracy, uncertainty, relationships to other variables, etc.). Variable names should be informative, short and, again, not include spaces or special characters. Here too, one should identify if and where there are missing data, and attribute the reason for missing data appropriately (e.g. was there a sample but no response, or was there no sample taken?). It is important to distinguish between no data and zeros. Where there are no data, the code "NA" should be used, instead of a zero value or a blank record.

8.4 Other key considerations

Modelling techniques such as ABMs and BNs often need data that are, in some way, manipulated (i.e. standardised scores, logarithmic scales, categorised data, etc.). It is therefore important that data are received in the most "raw" format as possible. Data should not be summarised or categorised before being shared (or, if this is necessary, it must be fully described in the "readme.txt" file). For example, if the absolute age of a person is recorded during a survey, then it is better that this data is provided in full, and reduced at a later stage, than recording data in pre-determined categories. Indeed, once data of this sort are lost, they cannot be recovered. If continuous data need to be categorised, then it is imperative that the same categories are used across all REFRESH sub-projects. Nested data should be listed in adjacent columns in the dataset, and clearly indicated as such in the metadata. Data should be ordered in columns with the first columns being sample indexes, the next set of columns being any nesting, followed by the outcome measures, the interventions or impacts (where relevant) and, finally, the covariates.

Box 1. Take outs – Data protocol.

Naming datasets and formats:

- Save data as ".txt" or ".csv" only.

- Keep dataset names to a minimum length (e.g. "Cons_beh_BBN_130616.csv").

- Do not use special characters (@, ~, *, etc.) in the dataset name.

- Include a version number or date in the dataset name (e.g. "Bus_beh_ABM_v8.txt").

- Do not include spaces in the name of the dataset – replace spaces with an underscore.

Dataset description and metadata:

- Save data description and metadata in a separate file named "readme.txt".

- Include a brief description of the dataset and how it was collected or collated.

- Choose short and descriptive names for variables, without spaces, capitalisation or special characters.

- Define variables fully; giving the units the data are in as well as any other relevant information (sources of data etc.).

- Identify missing data with the code "NA" to distinguish them from zero values – do not leave any blank cells in the dataset.

Other data considerations

- Include all the raw data – do not summarise data and convert continuous data in to categorical data.

Box 2. Example of metadata for sharing between work packages.

Datasets should be shared with both the data and associated metadata. Here an example with the dataset name of "EUBar388_010116.csv" is used.

An example dataset.

serialid	isocntry	q9	q1	q2_1	Age
1	Belgium	5% or less	Somewhat important	Somewhat positive	62
2	Belgium	5% or less	Very important	Somewhat positive	50
3	Belgium	5% or less	Somewhat important	Somewhat positive	40

Metadata should be saved in a text file (called Readme.txt) according to the criteria below.

Metadata:

- EUBar388_010116.csv – contains data from the Eurobarometer Flash survey 388 "Attitudes of Europeans to waste management and resource efficiency". The surveys took place in December 2013 and published in June 2014 (see http://ec.europa.eu/public_opinion/flash/fl_388_en.pdf). Flash Eurobarometers are ad hoc thematic telephone interviews allowing rapid data collection. Reproduction is authorised, except for commercial purposes, provided the source is acknowledged.

Citation(s):

- Report: European Commission, 2014a. Attitudes of European Towards Waste Management and Resource Efficiency. Report. Flash Eurobarometer survey n. 388, June 2014. Eurobarometer.

- Data: European Commission, 2014b. Flash Eurobarometer 388 (Attitudes of Europeans Towards Waste Management and Resource Efficiency). TNS Political & Social [producer]. GESIS Data Archive, Cologne. ZA5896 Data file Version 1.0.0. http://dx.doi.org/10.4232/1.11994.

Variable names:

- "serialid" – The sequential identifier for each record.

- "isocntry" – The name of the country where the respondent was questioned.

- "q9" – Question 9 asked respondents – "Can you estimate what percentage of the food you buy goes to waste?" The potential answers were categorical: 5% or less, 6% to 15%, 16% to 30%, 31% to 50%, More than 50%.

- "q1" – Question 1 asked "How important is it for you that Europe uses its resources more efficiently?" The potential answers were categorical: Did not know or answer (abbreviated to DK/NA), Not important at all, Not really important, Somewhat important, Very important.

- "q2_1" – Question 2 asked a multiple level question "In your opinion, what would be the impact of more efficient resource use on economic growth in [OUR COUNTRY]? The potential answers were categorical: Did not know or answer (abbreviated to DK/NA), Somewhat negative, Somewhat positive, Very negative, Very positive.

- "Age" – The age of the respondent. The data are continuous.

9 References

Anderson, P.W. 1972. More is different. *Science*. 177 (4047): 393-96.

Armstrong, M., and S. Huck. 2010. Behavioural Economics as Applied to Firms: A Primer. *Competition Policy International*. 6 (1): 1-45.

DellaVigna, S. 2009. Psychology and Economics: Evidence from the Field. *Journal of Economic Literature*. 47 (2): 315-72.

Digal, L.N., and F.Z. Ahmadi-Esfahan. 2002. Market power analysis in the retail food industry: a survey of methods. *Australian Journal of Agricultural and Resource Economics*. 46 (4): 559–84.

Di Maggio, P.J., and W.W. Powell. 1983. The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields. *American Sociological Review* 48 (2): 147-60.

European Commission 2014. *The economic impact of modern retail on choice and innovation in the EU food sector. Final Report*. Brussels: European Union.

Farmer, J.D., and D. Foley. 2009. The economy needs agent-based modelling. *Nature*. 460 (7256): 685-86.

Fehr, E., and U. Fischbacher. 2003. The nature of human altruism. *Nature* 425 (6960): 785-91.

Financial Times Lexicon. 2016. http://lexicon.ft.com/. Accessed on 3 September 2016.

Gallegati, M., and A. Kirman. 1999. *Beyond the representative agent*. Cheltenham: Edward Elgar Publishing.

Gallegati, M., and A. Kirman. 2015. *20 years of WEHIA a journey in search of a safer road*.

Gilbert, G.N. 2008. *Agent-based models*. Thousand Oaks (Ca): Sage Publishing.

Gode, D.K., and S. Shyam. 1993. Allocative efficiency of markets with zero-intelligence traders: Market as a partial substitute for individual rationality. *Journal of political economy*. 101 (1): 119-37.

Hakansson, N.H., J.G. Kunkel, and J.A. Ohlson. 1982. Sufficient and necessary conditions for information to have social value in pure exchange. *Journal of Finance*. 37 (5): 1169-81.

Haruvy, E., Y. Lahav, and C.N. Noussair. 2007. Traders' Expectations in Asset Markets: Experimental Evidence." *American Economic Review*. 97 (5): 1901-20.

Hatak, I., and K. Hysop. 2015. Cooperation between family businesses of different size: A case study. *Journal of Co-operative Organization and management* 3: 52-9.

Holland, J.H. 2000. *Emergence: From chaos to order*. Oxford: Oxford University Press.

Jackson, M.O. 2008. *Social and economic networks*. Vol. 3. Princeton: Princeton University Press.

Kahneman, D. 2011. *Thinking, fast and slow*. New York: Farrar, Straus and Giroux.

Kirman, A.P. 1992. Whom or What Does the Representative Individual Represent? *Journal of Economic Perspectives*. 6 (2): 117-36.

Verda, K. and Dragicevic, S.R. 2013. Bayesian networks and agent-based modelling approach for urban land-use and population density change: a BNAS model. *Journal of Geographical Systems*, 15: 403–426. Mas-Colell, A., M.D. Whinston, and J.R. Green. 1995. *Microeconomic theory*. Vol. 1. New York: Oxford University Press.

May, K. 1946. The aggregation problem for a one-industry model. *Econometrica: Journal of the Econometric Society*. 14 (4): 285-98.

Meyer, J.P., and N.J. Allen. 1991. A three-component conceptualization of organizational commitment. *Human Resource Management Review* 1 (1): 61-89.

Milgrom, P., and N. Stokey. 1982. Information, trade and common knowledge. *Journal of Economic Theory*. 26 (1): 17-27.

Östergren, K., J. Gustavsson, H. Bos-Brouwers, T. Timmermans, O.-J. Hansen, H. Møller, G. Anderson, C. O'Connor, H. Soethoudt, T. Quested, S. Easteal, A. Politano, C. Bellettato, M. Canali, L. Falasconi, S. Gaiani, M. Vittuari, F. Schneider, G. Moates, K. Waldron, and B. Redlingshöfer. 2014. FUSIONS Definitional Framework for Food Waste. Wageningen: FUSIONS Project. 134 pp.

Parfitt, J., M. Barthel, and S. Macnaughton. 2010. Food waste within food supply chains: quantification and potential for change to 2050. *Philosophical Transcripts of the Royal Society B: Biological Sciences*. 365 (1554): 3065–81.

Parsons, T. 1968. On the Concept of Value-Commitments. *Sociological Inquiry* 38 (2): 135-60.

Perloff, J. 2008. *Microeconomics Theory & Applications with Calculus*. Pearson.

Pope, A., and R. Gimmlet. 2015. Linking Bayesian and agent-based models to simulate complex social-ecological systems in semi-arid regions. *Frontiers in Environmental Science* (http://dx.doi.org/10.3389/fenvs.2015.00055).

Priefer, C., J. Jörissen, K.R. Bräutigam. 2013. *Technology Options for Feeding 10 billion People. Options for Cutting Food Waste. Study*. Brussels: European Union.

REFRESH Deliverable 4.1a. 2016. *Socio-economic implications of food waste: Consumers behavioural economic interrelationships and typologies* (to be published).

REFRESH Deliverable 4.1b. 2016. *Socio-economic implications of food waste: Business behavioural typologies and interrelationships* (to be published).

REFRESH Deliverable 4.1c. 2016. *Socio-economic implications of food waste: Economics of innovation* (to be published).

Rokeach, M. 1973. *The Nature of Human Values*. New York: Free Press. 438 pp.

Rubinstein, M. 1975. Securities Market Efficiency in an Arrow-Debreu Economy. *American Economic Review*. 65 (5): 812-24.

Schwartz, S.H. 1992. Universals in the content and structures of values: Theoretical advances and empirical tests in 20 countries. In M. Zanna (ed. by). *Advances in experimental psychology* Vol. 25. Orlando, FL: Academic Press. 1-65.

Sexton, R.J. 2013. Market Power, Misconceptions, and Modern Agricultural Markets, *American Journal of Agricultural Economics*. 95 (2): 209-19

Simon, A.H. 1956. Rational choice and the structure of the environment. *Psychological review*. 63 (2): 129-38.

Simon, A.H. 1957. *Models of Man: Social and Rational.* New York: John Wiley and Sons, Inc.

Stenmarck, Å., C. Jensen, T.E. Quested, and G. Moates. 2016. *Estimates of European food waste levels*. Wageningen: FUSIONS Project.

Sylos Labini, P. 1957. *Oligopolio e progresso tecnico*. Milano: Giuffrè Editore.

The New Palgrave Dictionary of Economics. 2013. http://www.dictionaryofeconomi cs.com/dictionary. Accessed 3 September 2016.

Tostivint, C., K. Östergren, T. Quested, H. Soethoudt, Å. Stenmarck, E. Svanes, and C. O'Connor. 2016. *Food waste quantification manual to monitor food waste amounts and progression*. Wageningen: FUSIONS Project. 165 pp.

Tversky, A., and D. Kahneman. 1973. Availability: A heuristic for judging frequency and probability. *Cognitive psychology* 5 (2): 207-32.

Vlaholias, E., K. Thompson, D. Every, and D. Dawson. 2015. Charity Starts… at Work? Conceptual Foundations for Research with Businesses that Donate to Food Redistribution Organisations. *Sustainability* 7: 7997-8021.

Von Neumann, J., and O. Morgenstern. 1953. *Theory of Games and Economic Behavior*. Third Edition. Princeton: Princeton University Press.

Wolfram, S. 2002. *A new kind of science*. Vol. 5. Champaign, UK: Wolfram Media.

10   Appendix

10.1 Glossary of socio-economic terms

Acyclic (*specific to BN*): Characteristic of BNs according to which it is not possible to have feedback loops in the model: causality occurs in a single direction.

Agent (or *individual*, or *entity*): Economic entity (business or consumer) who acts to achieve its purposes, using heuristic or simple decision-making rules, experiencing adaptation and interaction, and whose behaviour is modelled within agent-based models.

Agent-based model (*ABM*): Class of computational probabilistic models for simulating the actions and interactions of autonomous agents (both individual and collective entities, such as organizations or groups), with a view to identifying emerging issues at different scales, and assessing their effects on the system as a whole.

Alliance: Agreement, formal or informal, between two or more businesses, in order to perform activities together; it differs from a network since its participants are connected to each other by a contract, or by a set of more or less formalized rules.

Altruism: Attitude of individuals who perform «costly acts that confer economic benefits on other individuals» (Fehr and Fischbacher 2003, 785).

Bayesian Network (*BN*): Probabilistic graphical models which are used for representing a set of random variables and their conditional dependencies by means of directed acyclic graphs (without univocal direction of the relations among nodes); they allow scholars to analyse decision problems under uncertainty.

Behavioural economics: Science that studies the effects of psychological, social, cognitive, and emotional factors on the economic decisions of individuals and organizations, and the consequences for market conditions and resource allocation.

Bounded Rationality: Characteristic of economic agents with limited computational capacities, whose decision-making process is constrained by the way they process information.

Business (or *firm*, or *company*): Privately owned organization involved in the provision of goods, services, or both to consumers (final or intermediate), in exchange for other goods, services, or money. It refers to a particular organization, such as a single processing company or a large-scale retailer, and not to the entire market sector; within this report, it is intended to include food processors and retailers.

Cluster: «Concentrated group of companies in a particular field, [whose] activities are closely related and complementary to each other [...]. The companies involved include upstream and downstream buyers, as well as suppliers and producers» (Financial Times).

Commitment: Set of «moral obligations of the units of a system of social interaction to maintain the integrity of a value-pattern and to strive toward its implemen-

tation in action through combination with non-value factors» (Parsons 1968, 135). Organizational commitment (of managers and employees towards an organization) has three components: affective, continuation, and normative (Meyer and Allen 1991).

Company: see *Business*.

Complexity (*specific to ABM*): Within an ABM, it indicates a larger variability of the outcomes in terms of standard deviations of the target variables, e.g. due to the larger number of covariates and interaction rules.

Conditional probability: In probability theory, it indicates a measure of the probability of an event to occur, given that another event has occurred.

Consumer: Individual (physical person) who pays for using services and/or commodities, and is the final user of goods (e.g., food), or services (e.g., restoration services) produced, distributed, and sold by businesses.

Emergent phenomenon (*specific to ABM*): A phenomenon that becomes evident as consequence of the interaction among many agents and that cannot be immediately reduced to the characteristics of the local interactions (Anderson 1972).

Entity: see *Agent*.

Environment and *institutions* (*specific to ABM*): Within an ABM, it indicates all external constraints and characteristics that influence all agents (or groups of them) and their interactions. In practice, it represents the set of all policies, taxes, laws, institutions, and characteristics of the natural environment.

Firm: see *Business*.

Food processor (*processing sector*): Economic agent (business) that processes raw materials to make food products, implementing one or more procedures. Different processors perform different types of activities for transforming raw materials into food products; these processes may be relatively simple (e.g. cleaning fruits), or more elaborated (e.g. manufactured of prepared meals).

Food retailer (*retail sector*): Business operator of the branch *wholesale, retail and markets* of the food supply chain that sells food to final consumers, differently from wholesalers, who sell goods to other businesses. It may refer to both modern grocery retail (hypermarkets, supermarkets, discount shops) and other forms of retail (independent and traditional shops, and "new modern retail").

Food waste: «Fraction of food and inedible parts of food removed from the food supply chain in order to be recovered or disposed (including: composted, crops ploughed in/not harvested, anaerobic digestion, bio-energy production, co-generation, incineration, disposal to sewer, landfill or fish discarded to sea)» (FU-SIONS definition).

Heterogeneous Agents Model (*specific to ABM*): An ABM characterized by the presence of many agents with different specifications of their characteristics (attributes, rules of behaviour, memory, and perception of the environment).

Heuristic (*rule of thumb*): «Simple procedure that helps find adequate, though often imperfect, answers to difficult questions» (Kahneman 2011, 98). It is frequently used when the decision-making process is affected by cognitive biases. It may lead

either to good outcomes, or to erroneous judgments. Two examples are assessing the likelihood of an event according to the easiness of recalling a similar one (salience, Tversky and Kahneman 1973), or retaining a previous decision in the face of new information. Another simple form of heuristic is imitation (Di Maggio and Powell 1983).

Individual: see *Agent*.

Interaction (*ABMs and BNs*): One of the three different approaches that can be used to integrate the inputs and the outputs of ABMs and BNs, in order to maximize the amount of information on consumers' and businesses' food waste produced and analysed, given the limited availability of data.

Interaction rule (*specific to ABM*): Within an ABM, they represent the constraints on how agents can interact; they can be represented in game theoretical form, or as economic exchanges, or as exchanges of information. Since not all interactions are always possible, exchanges typically happen on a defined *interaction space*.

Interrelationship: Exchange of opinions, imitation, coordination schemes, etc., that take place in the framework of networks, alliances and clusters.

Investment: Capital formation, i.e. «the acquisition or creation of resources to be used in production» (The new Palgrave Dictionary of Economics).

Isolation (*Behaviour*): Assumption of the neoclassical economic theory according to which representative agents make their decisions independently from each other, i.e. are not influenced in their decisions by the decisions of other agents.

Machine Learning: Subfield of computer science which gives computers the ability to learn without being explicitly programmed; it explores the construction of algorithms which can learn from, and perform predictive analysis on data.

Memory (*specific to ABM*): Within ABMs, it indicates the possibility for an agent of recalling past actions and interactions, as well as their results, thus representing a mechanism to condition the behaviour of the agent on past actions.

Monte-Carlo Simulation: Class of computational algorithms which rely on repeated random sampling to obtain numerical results; within ABMs, this method consists in choosing a given parameter set, and iterating the dynamics of the model many times, with different sequences of random numbers, in order to obtain statistics of the final outcomes (means, standard deviations, distributions).

Network (*specific to ABM*): Two or more economic agents (not necessarily operating in the same sectors) that are connected or work together, making use of meetings among each other in order to share information (especially on innovation), assist each other, etc. Differently from alliances, networks do not imply direct connections or collaborations among all partners.

Object-oriented Approach (*specific to BN*): An Object-Oriented Network is a BN that, in addition to the usual nodes, contains a node that represents an instance (outcome or node) from another network or another type of model.

Pro-social behaviour: Acting kindly and helpfully towards strangers also if caregiving is not part of one's professional role (Vlaholias *et al.* 2005). It includes philanthropic and altruistic behaviours.

Rationality (*Behaviour*): Assumption of the classical economic theory according to which the preferences of representative agents are complete (an agent can always choose which of two alternatives he considers preferable, or if neither is preferred to the other), and transitive (if A is preferred over B and B is preferred over C, then A is preferred over C). Rational agents are assumed to take into account all available information, all probabilities of events, and potential costs and benefits while making a choice, and to act consistently.

Representative Agent: The typical decision-maker (e.g. a consumer, or a firm) in a neoclassical economic model, where all agents are identical, or differ among each other, but act in such a way that the sum of their individual decisions is mathematically equivalent to the decision of one individual or of many identical individuals.

Selfishness (*Behaviour*): Assumption of the neoclassical economic theory according to which representative agents maximize their individual profit (well-being) without being influenced by the consequence of their actions on the wellbeing of others.

Trust: «Having sufficient levels of positive expectations regarding [a] partner's behaviour to feel able to commit valuable resources (e.g., financial, know-how, etc.) to the cooperation with that partner, despite the risk that [he] might take (unfair) advantage of this relationship, and abuse this trust» (Roessl 2005, cited by Hatak and Hyslop 2015, 6).

Typology: Specific psychological factor, identified by behavioural economics, that may potentially affect (either as a driver, or as a barrier) the adoption of technological innovations.

Uncertainty: A situation is said to involve uncertainty if the randomness faced by economic agents presents itself in the form of alternative possible events but there are no statistics available and, therefore, no calculable probabilities of the outcomes of alternative decisions.

Values: «Desirable goals, varying in importance, that serve as guiding principles in people's lives» (Schwartz 1992, 21); «enduring beliefs that a specific mode, or conduct, or end-state of existence is personally or socially preferable to an opposite or converse mode of conduct, or end-state of existence» (Rokeach 1973, 5). They may represent idiosyncratic characteristics of individuals, but also of businesses, or of the whole society.

Zero-intelligence Agent (*specific to ABM*): Within ABMs, it represents the simplest type of agent, who acts randomly and is therefore used as a benchmark (Gode and Shyam 1993).