Multi-criteria assessment of ethical aspects in fresh tomato systems: Plant genomics technology innovation and food policy uses

Michalopoulos T.[1,2], Oude Lansink A. [1], Heuvelink E. [3] and Hogeveen H. [1]

[1] Wageningen University, Business Economics Group, Wageningen, The Netherlands
[2] Wageningen University, Applied Philosophy Group, Wageningen, The Netherlands
[3] Wageningen University, Horticultural Production Chains, Wageningen, The Netherlands

Abstract— **Product assessment for imperceptible characteristics like environmental impact, healthfulness, naturalness, and fairness is a helpful tool in product innovation and for enhancing socially responsible conduct.**

In this study we apply multiple criteria analysis for the assessment of fresh tomatoes in terms of consumer perceptions regarding the above characteristics. The generated indices provide an explicit and comprehensive representation of consumer perceptions. Existing tomato products from the Dutch market are ranked alongside (reasonable conjectures of) potential products to be developed with the use of plant genomics technology.

The results are interpreted to provide insights into the socially optimal use of (plant genomics) technology for fresh tomato production. Policy uses are highligted.

Keywords— **Ethical assessment, corporate societal responsibility, multiple criteria.**

I. INTRODUCTION

Three main categories of attributes affect consumer demand for (food) products. Price, narrow use value for the consumer, and performance for issues like its impact on the commons or on nature, on future generations, and on distant people which the consumer does not expect to ever meet. Somewhat arbitrarily, but in accordance with a common use of the word, we will refer to this third category of issues as 'ethical'. A significant feature of ethical attributes is that they are usually imperceptible. That is, unless valid and credible information is available, the performance of products for these issues cannot be assessed.

Sufficient numbers of consumers nowadays express ethical concerns to allow sizeable values-based label markets survive at non-competitive prices. However, ethical issues still lag behind attributes like price and taste, in determining market demand. Nevertheless, the importance of relevant issues is not underestimated, not even outside niche markets. Indeed, acknowledging the significance of ethical issues, a wide variety of conventional producers declares voluntary commitments to costly codes of ethical conduct, spending money to communicate 'green' or 'humane' socially responsible profiles. Moreover, a range of participatory methods is nowadays employed to increase the chances that costly investments in novel technologies will result to socially acceptable products.

In this paper we present and discuss a method for the assessment of (food) products on the basis of imperceptible attributes of concern to consumers. A multiple criteria approach [1] is used to rank tomato systems' relative performance for selected attributes. The focus is on the potential acceptability of different genomics[1]-enabled fresh

[1] In the Dutch context plant genomics is described as a 'high-throughput technology' (allowing fast and massively analysis of genomes), focused on 'research by means of a large scale characterization of food products into the elucidation of the way genes, RNA, proteins and metabolites interact in the functioning of cells, tissues, organs and the complete organism and its environment, both in an individual or in populations of species, as well as between species' [2]. Insights from plant genomics research may be applied for the efficient development of novel plant varieties in two distinct ways, either through conventional plant breeding or through genetic modification.
Conventional applications use techniques like marker-assisted-breeding (MAB) to increase the efficiency of traditional breeding in exploiting the potential that is already present in the genetic diversity within a species [3,4]. MAB follows much of the process of traditional breeding for the generation of a large numbers of cultivars, except from that selection of the promising varieties happens with the help of genetic analysis in the laboratory and not according to the opinion of the breeder in the field. Genetic modification techniques may be divided into cisgenic and transgenic. Like MAB, cis-genesis is restricted to the use of genetic diversity within a species. Unlike MAB, cis-genesis does not use the traditional plant breeding process, but it makes use of more efficient gene insertion techniques. Trans-genesis uses gene insertion techniques to introduce desirable genes from outside the genetic pool of a species. As such, transgenic techniques bear the highest positive potential and, as it has often been contested, the highest risks. The positive societal and economic potential of plant genomics applications consists of the development of plant varieties with improved nutritional value, resistance to pests, yield to inputs, tolerance of unfavourable environments, colour, taste, aroma, etc [e.g. 5, 6]. Criticisms of plant genomics mainly focus on the naturalness of cisgenic and transgenic applications [7], and on possible environmental and health risks. Further concerns

tomato scenarios. Straightforward uses for the indices would include informing consumer choice on imperceptible food characteristics of concern (like impact on climate change), assessing the societal performance of producers, advertising superiority over a competing product, and anticipating the societal acceptability of different trajectories for (plant-genomics) technology development.

However, this kind of assessment demands that a number of significant decisions are taken, chiefly what are the issues in terms of which the assessment is to proceed, and what is their relative importance. What needs to be determined here is not less than 'what matters' and 'how much' among the implications of (food) production, often hardly a rational puzzle to solve. To address such value judgments Michalopoulos *et al*. [9] proposed the so-termed 'ECHO Framework' for the 'ethical characterization' of foods.

II. MATERIALS AND METHODS

A. Assessment Criteria and Importance weights

Within the ECHO framework assessments proceed in two steps. First, 'structural input' in the form of assessment criteria and relative weights is collected. Structural input is meant to represent perceptions of citizens or consumers (depending on the purpose of the assessment) regarding '*what* matters and *how much*' in food production. The second step concerns the collection of input about the performance of products for the selected criteria ('product input'), and the use of multiple criteria modelling for the generation of concise product rankings.

An overall schematic representation of the assessment criteria used in this exercise is presented at the value tree of Figure 1. Their relative importance and the indicators used for their measurement are presented at Table 1.[2] Foods are assessed on the basis of four characteristics: Environmental Impact, Healthfulness, Naturalness, and Fairness of production. These characteristics were selected because they refer to *credence* food attributes (that is to attributes that consumers cannot assess either with their senses, or on the basis of previous consumption). Also, on the basis of their wide acceptability these characteristics may be *assumed* to be are *reasonable* (i.e. in agreement with the principles of justice and able to be described as parts of scientifically irrefutable views of the good life, ibid). Finally, the selected characteristics are *pragmatic* in the sense that they are important to an economically significant number of consumers. This is demonstrated in the market by the survival of organic, biological, fairness, and health labels at usually non-competitive prices. In particular, the four assessment characteristics were conceived so as to include the aspiration kernels of typical 'healthy' (or 'functional', 'fortified'), 'biological', and 'fair trade' food consumers, while the 'naturalness' category was meant to capture those aspirations of 'organic' food consumers that are not covered by 'health', 'biological', and 'fair trade' labels.

Each characteristic was described using criteria that are relevant to tomato production and to the options included to this assessment. The environmental assessment regards traditional environmental issues like water use, generation of wastes, impact on biodiversity, and also the timely concern of greenhouse gasses release. Impact on biodiversity included impact on non-target organisms, referring to the 'silent spring' effect [10], conversion of natural habitats to agricultural land, and perceived uncertainty about the environmental impact of (certified) GMO's. The Health category considers added (certified) benefits to consumer health about serious widespread diseases (such as cancers and cardiovascular diseases), health threats like pesticide residuals and pathogenic microbes, and also perceived uncertainties regarding the health impact of (certified) GMO's. Food Naturalness considers the breeding approach, the use of agrochemicals and of artificial infrastructure during farming, and the rooting environment of production plants (e.g. soil or artificial substrate). To assess Fairness, wealth distribution was considered alongside other publicly debated issues like the socially acceptable management of the commons and impact on global hunger. Wealth distribution regards the relative weakness of the producing economy and income of the 'weakest link' in the supply chain (farm labour, in this application). Democratic decision-making refers to the management of the commons, including the acceptability of irreversible changes and the commercial use of common resources. Impact on global hunger is assessed in terms of changes in the volume, the distribution, and the security of food production (minimizing threats).

regard the impact on the distribution of power within the agro-food sector and on farmers' autonomy from the protection of intellectual property rights (IPR) on genomics-enabled varieties [6, 8].

[2] To select criteria and indicators for assessing tomato products for each characteristic, relevant consumer concerns were mapped with the help of literature research. Further insight was obtained from an interdisciplinary food expert workshop. During the workshop a wide variety of experts from social and natural sciences, the production sector, and the civil society were invited to brainstorm on concerns relevant to plant genomics and tomato production. The generated set of criteria and indicators for their measurement was structured, complemented and refined using semi-qualitative interviews with disciplinary food experts. Next, written questionnaires were used to weight the relative importance of the selected criteria on the basis of Dutch consumer perceptions (N=101; unpublished data).

The performance of a selected set of options for these criteria was estimated on the basis of literature research and expert judgment. Next, products' performance for different criteria was overall comparatively assessed using a multiple-criteria method by Diaz-Balteiro and Romero (2004) [1]. These steps are discussed in more detail in the sections that follow.

HERE PLACE FIGURE 1

HERE PLACE TABLE 1

B. Tomato Options and Performance Matrix

Eleven fresh tomato options were considered in this exercise (Table 2). The assessed tomato options fall under two categories. The first category includes existing tomato 'systems'. 'Systems' comprise the end-product (fresh tomato) and aspects of their production history that are relevant to the selected criteria. The second category comprises currently experimental or hypothetical scenarios of tomato systems (explained below). Only systems with significantly different performance for at least one of the selected criteria were included in this exercise. Also, because focus here is on imperceptible and ethical characteristics, tomatoes' differences in price, colour, etc. were out of scope and not considered.

The selection of tomato systems was meant to provide insight on the optimal embedment of novel technology (plant genomics) in food production.

HERE PLACE TABLE 2

The first category (options O1 to O4) comprises existing tomatoes that can be found in the Dutch market. Existing tomato options may be described with different degree of detail, depending on the purpose of the study, and also depending on data availability. For the purpose of the present exercise it was thought sufficient to use a rather high level of abstraction, specifying five broad categories of tomato systems as presented at Table 2. These options include so-termed 'standard' and 'organic' systems from The Netherlands and from Spain[3]. The 'standard' options refer to (for all relevant purposes and for all relevant indicators) stereotypic, mainstream, or bulk tomatoes[4]. The 'organic' options refer to representative tomatoes with nationally recognized organic certifications. The options included in this category were scored on the performance matrix according to their average performances for each criterion. These performances were estimated for the purpose of this illustration, from published sources [12, 13, 14, 15, 16, 17, 18], personal communication with experts, and with consumer organizations. Considered assumptions were made when needed data was hard to mine. Hypothetical product scenarios were assumed. All product input is shown at the Performance Matrix (Table 3).

The second category (options O5 to O11) comprises imaginary or potential systems. System O5 (Advanced-NL) represents the fraction of Dutch tomatoes that are produced in technologically advanced ('semi-closed') glasshouses [19]. The FairBasic-Ma option (O6) has been defined to represent an imaginary system that meets *only the minimal* Fair-Trade standard of 'a fair price', thus fulfilling the basic requirements of the Fair-Trade trademark. It has been introduced to serve as a fairness standard to compare with O9. The properties of this hypothetical system have been defined on the basis of data about Spanish and Moroccan production and practices and climatic conditions. Morocco has been chosen as the land of production so that the system satisfies the Fair-Trade focus of production in developing countries, and because Moroccan tomatoes may be found in the Dutch market. It has been assumed that the system delivers a traditional variety, neither MAB nor genetically modified. Environmental and healthfulness performance was scored according to regional norms. When data about the performance of Moroccan tomatoes was limited, then data for StandardES served for scoring this system.

Systems O7, O8, and O9 refer to non-genetically modified products, produced with the use of MAB technology. They have been purposefully defined to represent reasonably realistic scenarios of using MAB technology to apply insights from plant genomics research with the sole purpose to improve systems' environmental, healthfulness, and fairness characteristics, respectfully. Scoring these non-existent systems at the performance matrix was partly based on tomato and genomics experts'

[3] Spanish tomatoes in the Dutch Market are usually produced in Almeria and Murcia [11], and therefore data from those regions have been considered in this study. Dutch tomatoes originate mainly from Westland.

[4] The relevant purpose is to be distinguishable. A tomato may be regarded as 'standard' either when it does not perform significantly different from the bulk of production at its country of origin, or when in reality it does perform differently but there is lack of a credible traceability system to communicate that information (for more on 'ethical traceability' see [20]). That is, a system is assessed on the basis of information that is available. When information is not available, then the system is regarded to be standard. When that is not true and the system is –in reality- above standard then suppliers may inform for corrections. When a system is –in reality- below standard, then corrections may be expected by the civil society (NGO's). What systems are considered as standard affects the properties of the standard options.

opinion. MAB-enabled systems were assumed to deliver a 15% improvement for each genetic characteristic relevant. Besides, each of these options was conceived as a modification of O1, which means that when their performance for some relevant criterion would not be affected by MAB, then the options' scores would remain the same as for O1 (that is, O1 was used as a performance benchmark -a point of reference- for scoring these options regarding MAB-irrelevant criteria).

Systems O10 and O11 are genetically modified. Including these options in the assessment is meant to return an indication about the range within which genetically modified products could perform. To be fit for this purpose options O10 and O11 have been intentionally defined with in-built biases: Option GM-A-NL (O10) was conceived to represent a reasonably realistic in the mid-term scenario, in which genetic modification is used to bring about an overall improvement for the selected criteria. Consequently it has been defined to carry the most desirable (or the least controversial) attributes for each category of concerns. It is defined to be cisgenic so as to represent the least problematic genetically modified product scenario from a naturalness perspective [21]. Similarly, option GM-B-NL (O11) was conceived to represent a reasonably realistic scenario in which genetic modification is used to improve some characteristic that is irrelevant to the focus of the present study (for example, a tomato's taste or colour). Consequently O11 has been attributed the least desirable (or the most controversial) attributes for each category of concerns. No improvement for any relevant criterion was assumed as compared to O1, while worse performances were attributed when thought realistic, namely regarding yield decrease[5]. It is defined to be transgenic with inserted genes originating from a different realm so as to represent the most problematic genetically modified product scenario, from a naturalness perspective. Scoring non-existent systems O10 and O11 at the performance matrix was also based on expert opinion and relied on O1 as a point of reference. We assumed the risk perceptions of cisgenic systems to be the same as that of the transgenic (i.e. for GMO spread or for safety uncertainty (Healthfulness category), possibly an underestimation of the public perception of cisgenic products.

HERE PLACE TABLE 3

[5] A 10% yield decrease was assumed. A tomato's taste depends, among others, on the concentration of certain ingredients. In principle, decreasing the 'dilution' of such substances enhances the taste of the product. See Morris and Sands [22] for a discussion on unintentional worsening of certain characteristics as a side-effect of trying to improve others.

C. Empirical Model

The performance of different tomato options was ranked for each characteristic according to the multiple-criteria methodology presented by Diaz-Balteiro and Romero [1]. Multiple-criteria methodologies require the description of characteristics by means of a set of subjectively or objectively measurable indicators (criteria), which are related with relative importance weights. In spite of the underlying difficulties (like hidden nonlinearities, interaction between indicators, dynamic aspects, etc), multiple-criteria modelling offers a pragmatic approach to complex problems where no single consensual indicator is available for the assessment of different options, like in sustainability studies (e.g., see Ref Rennings and Wiggering [23], Pannell and Glenn [24]). In short, multiple criteria modelling is appropriate for the characterization of different (food) options because: it is theoretically sound [25], it can incorporate objective as well as subjective indicators, it can rank an unrestricted number of alternatives, it is simple to use, and it is transparent [26]. The selected methodology can be used to provide solutions for the problem of comparing the performance of different options as a whole for a set of specified criteria [1, 26].

To account for differences in measurement units and also to uniformly transform all indicators to the type "more is better", application must start with a normalization step. Performances for different indicators were normalized as follows [1]:

$$\overline{R}_{ij} = 1 - \frac{R_j^* - R_{ij}}{R_j^* - R_{*j}} = \frac{R_{ij} - R_{*j}}{R_j^* - R_{*j}} \forall i.j \qquad (1)$$

Where, for each characteristic i=1, 2,…,n are different options evaluated according to j=1,2,…,m indicators. R_j^* is the optimum value of the jth indicator (ideal value). This optimum value represents a maximum value if the indicator is of the type ''more is better'' or a minimum value when the indicator is of the type ''less is better''; R_{*j} is the worst value achieved by the jth indicator (anti-ideal or nadir value); $\overline{R}_{ij}$ is the normalized value achieved by the ith system with respect to the jth indicator. Using the above normalization system, the indicators have no dimension and are also bounded between 0 and 1. Hence, for this normalization system the ideal vector is $\overline{R}^*$=(1,. . .,1) and the anti-ideal vector is $\overline{R}$ =(0,. . .,0) (ibid). As seen before, these vectors were introduced in the comparison as 'Best' and 'Worst' options, for each characteristic.

For objectively measured criteria like GHG emissions, or pesticide residuals, we assumed that the best performance

was the most preferable ('more is better' or 'worse is better', depending on the criterion). For subjectively measured criteria with more than one level, like in the case of 'naturalness of creation', preference scales were constructed according to expert opinion. To relax data requirements, local scales have been used for the normalizations.

Next, the different options were ranked for each characteristic according to the function:

$$S_i = (1-\lambda)[\underset{j}{Min}(W_j \overline{R}_{ij})] + \lambda \sum_{j=1}^{m} W_j \overline{R}_{ij} \quad (2)$$ [6]

Where, S_i is the generic i[th] food option; *ij* is the generic j[th] indicator of the characteristic; *Wj* is the weight or relative importance attached by an expert or by a panel of experts to the jth indicator; and λ is a control parameter that takes values $0<\lambda<1$. The control parameter λ in Eq.(2) measures the tradeoffs between the 'aggregated' and 'most balanced' rankings. When λ=1, then the model assumes independence among the specified indicators, and additive aggregate rankings are produced. In this case the options are ranked for their total performance ('according to their weighted sum of the normalized indicators'). When λ=0 additivity is not assumed and so-termed 'most-balanced' rankings are produced. In this case the options are ranked according to their performance for the single indicator for which they perform the worst ('for the minimization of the maximum deviation of the most displaced indicator from the ideal'). For in between values of the control parameter λ solutions that correspond to 'different additivity levels' are obtained (ibid.).

III. RESULTS

The resulting comparative rankings of tomato systems are presented at Table 4. The first four sections present indices for each of the four characteristics, while the fifth section represents a possible overall assessment. Each score represents a systems' performance as compared to vectors of best and worst achieved values for each criterion, which perform the role of upper and lower boundaries, scoring 0 and 1, respectively. At the leftmost of each section are presented the rankings for the case that we accept complete additivity, namely the aggregated indices (λ=1,0). Next is presented a situation of low additivity (λ=0,1) and finally the most-balanced (λ=0,0) indices (all in between possibilities, referring to different degrees of additivity, may also be calculated). Aggregated assessments are quite usual in policy decision-making. The resulting indices assume that indicators are independent and that the represented values are commensurable. Commensurability implies that lower performances for one indicator can be compensated by higher performances for another. Consequently, it is possible that options with high aggregated rankings may score low (sometimes perhaps even unacceptably low) for certain assessment criteria. Such tradeoffs might not be always ethically acceptable [28], however they might be necessitated on pragmatic grounds [29]. The overall aggregated index presented at Table 4.e is an example of a possibly contestable aggregation[7] because it assumes that one may compare and substitute achievements of quite different values (i.e. to compare gains or losses for environmental impact to those of healthfulness)[8]. Notwithstanding the relevant debate, what tradeoffs are legitimate remains, essentially, a political decision to be taken on a case-by-case basis.

When tradeoffs are not welcome the most balanced index compares options for their single worst performance. This has the consequence that whenever an option scores the worst (anti-ideal) performance for at least one indicator, then it shall also score zero at the most-balanced index. When more than one options score zero then it is possible to differentiate among them by comparing the part of the total importance weight for which options score the anti-ideal values. This ranking is presented at columns 'ZeroRank'. For instance, when ZeroRank=0,1 then this means that the option has the worst possible values for (1-

[6] This function is mathematically similar to the augmented Tchebycheff function, which has been widely researched in the MCDM literature [1, 30 chapters 14 and 15, 31].

[7] The overall aggregated assessment (and also possible further product optimization) is also demanding in that it expands the requirement for independence from indicators belonging to the same characteristic to indicators describing different product characteristics. When some trait is important from the perspective of different characteristics, like in the case of genetic modification, then this requirement may create computational difficulties.

[8] Besides, in the current exercise there are two case of depended variables across the different categories: 'Yield to land' is considered both for the Environmental (sufficient food production volumes reduce pressure to convert natural habitats to agricultural land) and the Fairness categories (increased food yields contribute to combat world hunger). Escape of GMO's is a potential risk to biodiversity (environment), and also may bring about irreversible changes in the commons (natural environment) which can be seen as undemocratic when public opinion disagrees with such changes. Such dependencies depend on the particular application of the method, and may be avoided with a different selection of indicators. The issue of commensurability between different values, however, touches upon a more fundamental theoretical issue that will be common to all cross category aggregations. Within the same concern tradeoffs might be more meaningful. Non aggregated results are also presented at the rightmost column of Table 4.

ZeroRank)*100%=90% of the total importance weight assigned to the assessment indicators[9]. Alternatively, a low degree of additivity may be tolerated (for example see ranking for λ=0,1). What level of additivity is assumed may influence the final rankings.

HERE PLACE TABLE 1

Interpretation of the results is not absolute because of the absence of (agreement on) absolute benchmarks needed for such assessments. Such benchmarks would mark what is acceptable (e.g. environmental) performance, and would allow one judge an option to be 'positive' or 'negative' (e.g. 'good vs. bad', 'environmentally friendly vs. unfriendly', 'natural vs. unnatural', 'fair vs. unfair', etc.). Agreement on such benchmarks often appears hard to achieve, and efforts to that direction result in counterproductive polarizations [32].

These difficulties may be circumvented by using relative referencing. For example, the StandardNL option may serve as a reference system to base interpretations of the performance of the other options. StandardNL[10] is by definition a good candidate for this purpose because it represents the typical, or 'bulk' of tomato to be found (which means consumed) in the Dutch market. This allows claims about an option's performance as compared to '*what is the norm*' locally. Consequently, options that score above StandardNL may claim to be improvements (superior) as-compared-to-the-norm-of-tomatoes-usually-met-in-the-Dutch-market, while an option that scores below StandardNL may be described as 'worsenings' (inferior).

IV. DISCUSSION

Accordingly, the aggregated indices of Table 4 compare the acceptability of the systems for a broad range of criteria and according to their perceived importance as stated by Dutch consumers.

Results for the four existing tomato systems (StandardNL, OrganicNL, StandardES, and OrganicES) are mostly intuitive. Organic Dutch production is leading for Environmental friendliness concerns achieving and overall 72% of the ideal performance. Organic Dutch tomatoes are followed by typical Dutch production, achieving a (1-0,62 =) 38% lower Environmental score. In spite of lower greenhouse gasses emissions, Spanish production's circumstances of water scarcity, higher agrochemical emissions, and lower yields make it lag overall behind the Dutch. Organic produce scores higher than the conventional one for Healthfulness concerns. Spanish production achieves a by more than 50% lower overall score, due to the higher chance that pesticide residues will be found in Spanish tomatoes at detectable levels. Organic systems also achieve the highest score for Naturalness, with typical Dutch production being the least natural. Typical Dutch production however best accommodates overall Fairness concerns due to its performance for world hunger (higher yield helps combat world hunger. Also, the lower the use of pesticides the lower the evolutionary pressure on pests to develop resistance and threaten food security).

Existing tomato system scores may be compared to hypothetical system scenarios. The scores achieved by AdvancedNL and by the MAB-enabled systems, as defined, indicate their potential to deliver improved achievements for all categories of concerns, with the exception of Naturalness. The technological improvements in greenhouse efficiency assumed for AdvancedNL allow it achieve the second best environmental performance, considerably narrowing to that of OrganicNL. The most promising improvements seem to be for the Healthfulness category, where the potential of MAB to deliver products with extraordinarily high content in certifiable functional health ingredients reveals a potential for MAB-Health to dominate organic production. The performance of the two genetically modified scenarios shows that, according to stated public perceptions, ethical acceptability varies significantly with the impact of genetic modification. This result confirms the need for case-by-case analysis of different research and development scenarios. As illustrated in the case of Environmental and Healthfulness concerns, systems delivering genetically modified products have the potential to be perceived as improvements to existing ones. Having said this, attention should be also drawn to the most-balanced rankings presented at the rightmost column of Table 4. The most-balanced rankings depict scores reached by the systems when tradeoffs between accomplishments of different criteria are not accepted. Consequently systems are scored on the basis of their worst performances for the criteria under consideration. In these cases, the worst scores consistently achieved by the modified varieties are indicative of their poor performance for a number of criteria. Exceptionally poor performances for certain criteria (even though compensated at the aggregated index) provide the factual support for arguments citing unacceptable drawbacks to reject genetic modification as a whole – despite the possible benefits. In the present application most

[9] Remember that results refer to system performances normalized so that 1 consistently represents the best outcome and 0 the worst. The best outcome in this case is that no indicator scores the anti-ideal value, while the worst outcome is that they all do.

[10] 'NL' as in 'StandardNL' refers to country of production, not consumption. However, in the case of The Netherlands and fresh tomato consumption, these coincide.

systems coincidentally score the worst value among the options for at least one indicator. These systems get attributed a most-balanced ranking of 0. For these reasons the presented most-balanced indices may be more meaningful as indicators of possible controversy about a system, than as performance rankings. For the reader that is interested in non-compensatory rankings, the included 'low aggregation' column might provide a useful approximation. The column 'ZERORANK' indicates the percentage of total importance weight attached to assessment criteria, for which the option does not score the anti-ideal score. That is, a low ZERORANK score indicates that from the perspective of a large part of respondents' concerns, the option does not seem to be a good idea.

Counterintuitive appear the consistently low rankings of the FairBasic-Ma system for almost every category of concerns, including that of Fairness. This is the result of the more inclusive understanding of fairness used by the present framework, than that of the Fair-Trade movement. Namely, democratic governance and world hunger issues are also taken onboard the fairness agenda besides wealth distribution. The performance of tomato systems for such criteria however may depend on the genetic properties of the cultivar (like yield and resistance), and on farming practices (pesticide use and irrigation efficiency). Strengthening certain environmental requirements of FairTrade production would improve the movement's ranking. Also counterintuitive could be considered that officially health-certified cisgenic products would be perceived to be more healthful than organic ones. Surely, such intriguing results could be further studied, especially since there are today examples of quite successful (at least within certain consumer groups) functional foods to be found in every supermarket. However, counterintuitive results may only indicate the need for model improvement. A clear candidate for such improvement is the crude perception of risk used in this application, in the form of weighted binary variables.

The meaningfulness of these statements depends on the input used for their generation. Input required for the present for the current framework may be distinguished to structural input and product input [9]. Structural input includes assessment criteria and attached importance weights. Product input includes options' performance for those criteria.

Product input requirements may vary with the application. For innovative producers, the model allows to assess different research scenarios, and also compare hypothetical product scenarios to existing competitors. Well performing actors may use results to improve allocation of research priorities, to defend against relevant criticism, and to illustrate the trustworthiness of promotion claims. Referring to national production averages for these comparisons has the advantage that data is often publicly available for a large number of major indicators. Besides, comparisons to average allow for a basic 'above average' (improvement), 'average' (business-as-usual), 'below average' (worsening) characterization. To get the assessments, the producer should have available a set of weighted criteria (structural input), on the basis of which to compare input on own products (scenarios) to national averages and competitor estimates. NGO's may use their estimations to criticize producers with poor performances. Possible uses for the government would include improving the allocation of research funds. Also, the rankings could provide the basis of product labelling, informing food consumer choice on some area of public concern with imperceptible aspects of production (perhaps the relative greenhouse gasses emissions of substitute products).

Comparing research scenarios to aid decisions on research agenda must demand high quality of input, and clear definition of what constitutes "reasonable expert conjectures". For the purpose of informing consumer choice however, product input requirements could be more relaxed. In particular, it would be besides the point to attempt top-down estimations of the actual performance of every production system. Systems that do not use separate supply chains could be aggregated. Resulting statistical averages of separate supply chains' performance would suffice for generating information that will be useful to actual consumer choice. Independent supply chains would have an interest to provide data that prove that they perform better than competitors. NGO's and consumer organizations could have a stake in revealing misreporting.

The possible economic impact of such uses indicates the importance of transparency while selecting the assessment criteria and attached importance weights. Michalopoulos et al. [9] proposed to periodically derive structural input from deliberative consensus workshops. In these workshops achievement on assessment criteria and attached importance weights should achieved by a representative, informed (by stakeholders), and deliberated public panel (citizens or consumers)[11]. This illustration remained true to the spirit of the suggestion by considering an inclusive account of concerns linked to food production, and by deriving importance weights from the public. Grounding decisions regarding value judgments such as 'what matters' in food production (and 'how much') onto public opinion instead of on that of experts, allows a claim for (inter-subjective) impartiality for the rankings delivered by the present assessment framework.

[11] That would amount to input about 'what matters' and 'how much' regarding food production.

V. CONCLUSIONS

Product assessment for imperceptible characteristics like environmental impact, healthfulness, naturalness, and fairness is a helpful tool in product innovation and for enhancing socially responsible conduct.

In this study we apply multiple criteria analysis for the assessment of fresh tomatoes in terms of consumer perceptions regarding the above characteristics. The generated indices provide an explicit representation of consumer perceptions that are usually reflected implicitly in consumer choices. Existing tomato products from the Dutch market are ranked alongside (reasonable conjectures of) potential products to be developed with the use of plant genomics technology.

The results are interpreted to provide insights into the socially optimal use of (plant genomics) technology for fresh tomato production. Policy uses for product differentiation and research agenda setting are highlighted.

ACKNOWLEDGMENT

This work has been conducted within a project studying the societal aspects of plant genomics for the Centre for Bio-Systems Genomics, The Netherlands. It has been also supervised by the Applied Philosophy Group of Wageningen University.

REFERENCES

1. Diaz-Balteiro, L, and Romero C, (2004) In search of a natural systems sustainability index Ecological Economics 49:401-405.
2. NWO (Netherlands Organisation for Scientific Research) (2001) De maatschappelijke component van het genomics onderzoek. Oproep tot het indienen van aanvragen. Den Haag: NWO, at www.nwo.nl/subsidiewijzer.nsf/pages/NWOP_547JJZ?Opendocument&su=n
3. Buckler IV E S and Thornsberry J M (2002) Plant molecular diversity and applications to genomics. Current Opinion in Plant Biology, 5, 107–111.
4. Mahalakshmi V, Ortiz R. (2001) Plant genomics and agriculture: From model organisms to crops, the role of data mining for gene discovery. Electronic Journal of Biotechnology at 4(2), http://www.ejbiotechnology.info/content/vol4/issue3/full/5/
5. Crusak M A, Cakmak I (2005). Methods to improve the crop delivery of minerals to humans and livestock. In M. Broadley and P. White, Plant Nutritional Genomics 277-280. Blackwell Publishing, Oxford
6. Nap J P, Jacobs J, Gremmen B, Stiekema W J (2002) Genomics and sustainability: exploring a societal norm, In NWO (Netherlands Organisation for Scientific Research) Essays. De maatschappelijke component van het genomics onderzoek. Den Haag: NWO
7. Gremmen, B (2005) Genomics and the intrinsic values of plants Genomics, Society and Policy 1:3:1-7
8. Calva E, Cardosa M, Gavilondo J, (2002) Avoiding the genomics divide. Trends in Biotechnology 20:368-370.
9. Michalopoulos T, Korthals M, Hogeveen H, (2008) 'Trading "Ethical Preferences" In the Market: Outline of a Politically Liberal Framework for the Ethical Characterization of Foods', Journal of Agricultural and Environmental Ethics 1
10. Carson R, (1962) Silent Spring, Hamilton
11. Costa J M, Heuvelink E (2005) Introduction: the tomato crop and industry, In: Heuvelink E (ed) Tomatoes, Crop Production Science in Horticulture Series 13, Oxford
12. de Haan B J F F P J, Sukkel .J en W (2007) Energieverbruik, broeikasgasemissies en koolstofopslag: De biologische en de gangbare landbouw vergeleken. PRI, Wageningen UR, rapport 140.
13. Van der Velden N J A, Janse, Kaarsemaker J R C, Maaswinkel R H M (2004) Duurzaamheid van vruchtgroenten in Spanje; Proeve van monitoring, LEI , Den Haag,
14. van Kooten O, Heuvelink E, Stanghellini C (2006) New developments in greenhouse technology can mitigate the water shortage problem of the 21st century. Acta Hort. (ISHS) 767:45-52 at http://www.actahort.org/books/767/767_2.htm
15. Costa J M, Heuvelink E 2000; greenhouse horticulture in almeria Spain
16. Pluimers J C (2001) An Environmental Systems Analysis of Greenhouse Horticulture in the Netherlands: the Tomato Case, PhD. thesis, Wageningen, The Netherlands, Wageningen University.
17. Muñoz P, Anton A, Montero J I, Castells F (2003): Using LCA for the Improvement of Waste Management in Greenhouse Tomato Production. DIAS report, Life Cycle Assessment in the Agri-food sector, Proceedings from the 4th International Conference, October 6-8, 2003, Bygholm, Denmark.
18. Pesticide residues monitoring second and fourth quarter report 2004 at http://www.pesticides.gov.uk/uploadedfiles/Web_Assets/PRC/PRC_2004_Q4_report.pdf

19. van der Ploeg A, van der Meer M, Heuvelink E (2007) Breeding for a more energy efficient greenhouse tomato: past and future perspectives Euphytica 158:129–138
20. Coff C, Barling D, Korthals M, Nielsen T, (Eds.) (2008) Ethical Traceability and Communicating Food.: The International Library of Environmental, Agricultural and Food Ethics Series, Vol. 15 ISBN: 978-1-4020-8523-9
21. Lammerts Van Bueren E T, Verhoog H, Tiemens-Hulscher M, Struik P C, Haring M A. (2007) Organic agriculture requires process rather than product evaluation of novel breeding techniques, NJAS, 54:4
22. Morris C E, Sands D C (2006) The Breeder's Dilemma—Yield Or Nutrition Nature Biotechnology 24 : 9
23. Rennings K, Wiggering H (1997) Steps towards indicators of sustainable development: linking economic and ecological concepts. Ecological Economics 20:25–36
24. Pannell D J, Glenn N A (2000) A framework for the economic evaluation and selection of sustainability indicators in agriculture. Ecological Economics 33:135–149.
25. Keeney R L, Raiffa H (1976) Decisions with multiple objectives: preference and value tradeoffs. Wiley, New York.
26. van Calker K J, (2005) Sustainability of Dutch dairy farming systems: A modelling approach, PhD Thesis,, Wageningen Universiteit, pp173
27. Hanley N, Moffatt I, Faichney R, Wilson M (1999) Measuring sustainability: a time series of alternative indicators for Scotland. Ecological Economics 28:55–73
28. Chang R (1997), ''Introduction,'' in R. Chang (ed.), Incommensurability: Incomparability and Practical Reason, USA: President and fellows of Harvard College, pp. 1–34.
29. Hardin G (1968), ''The Tragedy of the Commons.'' Science, 162:1243–1248.
30. Steuer R E (1989) Multiple Criteria Optimization: Theory, Computation, and Application. Krieger, Malabar.
31. Tamiz M, Jones D, Romero C (1998). Goal programming for decision making: an overview of the current state-of-the-art. European Journal of Operational Research 111:569– 581.
32. Keulartz J, Schermer M, Korthals M, Swierstra T, (2004) Ethics in Technological Culture: A Programmatic Proposal for a Pragmatist Approach, Science, Technology & Human Values 29:1:3-29