Socio-environmental multi-labelling and consumer willingness to pay

Ivan Dufeu, Jean-Marc Ferrandi, Patrick Gabriel, Marine Le Gall-Ely

Article

RAM

Recherche et Applications en Marketing
1–22

DOI: 10.1177/2051570714542063
rme.sagepub.com

SAGE

Socio-environmental multi-labelling and consumer willingness to pay

Ivan Dufeu
LUNAM Université – ONIRIS – LARGECIA Nantes, France

Jean-Marc Ferrandi
LUNAM Université – ONIRIS – LARGECIA Nantes, France

Patrick Gabriel
ICI EA2652, Université de Bretagne Occidentale, France

Marine Le Gall-Ely
ICI EA2652, Université de Bretagne Occidentale, France

Abstract

The objective of this article is to assess the impact of a gradual increase in the number of labels appearing on ~~a certain~~ food products on consumer valuation of the given product. Three empirical studies were designed to measure the effects of using labels to differentiate food products (Organic Farming, Fairtrade, and *Label Rouge* (a French label that concerns organoleptic quality)) on the willingness of 519 French consumers of honey to pay a premium. These three studies shed light on a complementarity effect, often dominating the effects of redundancy and information overload. This effect differed depending on the features and the number of associated labels, and was influenced by consumer trust in the labels. These findings should encourage producers to associate their products with labels on packaging when they are consistent in terms of reputation but complementary in terms of attributes.

Keywords

complementarity, labels, perception, substitutability, willingness to pay

Introduction

In recent years, there has been a significant increase in the number of logos on food packaging. These include both product or company brand logos, as well as 'collective' brand logos.[1] Among the latter, certification marks, more generally referred to as 'labels', have the particularity of being based on

Corresponding author:
Ivan Dufeu, LUNAM Université – ONIRIS – LARGECIA Nantes, Site de la Géraudière, CS 82225, 44322 Nantes Cedex 3, France.
Email: ivan.dufeu@oniris-nantes.fr

specifications published and monitored by independent bodies. These certification marks may be issued by public institutes, as is the case for the French agricultural labels Label Rouge (LR) and Agriculture Biologique (AB), as well as Protected Geographical Indication (PGI), or by private associations, such as Fairtrade/Max Havelaar (MH), for example.

In the case of food products, the application of collective certification marks (hereafter referred to as labels) plays a special role. It serves to inform and reassure consumers about various quality aspects (notably the product's credence attributes[2]) of the product that consumption alone cannot reveal. Beyond the additional information they provide, labels have the advantage of reinforcing the positioning and differentiation of the food product, which are particularly important factors in self-service distribution channels (Hamzaoui-Essoussi et al., 2013). The socially responsible labelling of a product is even likely to strengthen the position of not only the brand, but also the store selling this brand (Aouina Mejri and Bhatli, 2014). Finally, while labels tend to enhance product attractiveness and consumer trust, they also result in lower price sensitivity and an increased consumer willingness to pay (WTP) (Janssen and Hamm, 2012; Tagbata and Sirieix, 2010; Zander and Hamm, 2010). This WTP, or reservation price, is defined as the highest price a consumer is willing to pay for a product or service. It is the result of a positive attitude towards a product and its price (Le Gall-Ely, 2009).

The strong competition that exists in the agri-food sector and on supermarket shelves prompts many producers or industrial groups to accumulate labels. Their reasoning is that their brand image will be consolidated through increasing the benefits of labelling (Sirieix et al., 2013). Now it is common to see up to three labels printed side by side on food packaging. Alter Eco, the France-based alternative trading organization, has thus tended to accumulate quality marks on its products, including MH, AB, PwC audit, 'Produit compensé carbone' ('Carbon neutral product') and 'Agriculture Française Equitable' ('Sustainable French Agriculture'),[3] resulting in dense packaging. This multi-labelling can be further increased with the emergence of many public or private environmental labels which have significant marketing implications. However, the actual effects of combining several labels on food packaging are still relatively unexplored.

What marketing literature does confirm is that the implicit assumption of accumulated or increased benefits resulting from multiple labels is questionable. On the one hand, the effects of the labels on the consumer in terms of purchase intention or actual behaviour are not systematic. Rather, they depend on the perceived reputation of the label and of the individual brand with which it is associated (Larceneux, 2001; Larceneux, Benoit-Moreau and Renaudin, 2012; Janssen and Hamm, 2012). The WTP for the combination of a label and a brand varies, notably depending on brand image (Bauer et al., 2012; De Pelsmacker, Driesen and Rayp, 2005a). On the other hand, the presence of several labels can increase uncertainty and reveal various influences on consumer behaviour (Dekhili and Achabou, 2013; Janssen and Hamm, 2012; Tagbata and Sirieix, 2010). The combination of two labels must thus be perceived as relevant and complementary, otherwise it can erode the product value (Sirieix et al., 2013).

All in all, marketing research primarily highlights the effects produced by asymmetries between the messages conveyed by these different official labels or simple brand logos. Following on from this research, the purpose of the present study is to contribute to measuring the effects of the multi-labelling of food products where there is no apparent contradiction or asymmetry between the labels. Unlike previous research, examining labels with similar reputations and levels of consumer confidence, but with complementary specifications, makes it possible to focus on the effects of potential synergies related to the content of these specifications.

This research question meets real managerial concerns, as the multiplication of controlled labels of varying natures (and therefore requirements) implies substantial investment in order to comply with the specifications. Uncertainties about the marketing impact of multi-labelling explain the various industrial behaviours observed in this regard. Indeed, it seems that firms in the agro-food sector need more information on an effective policy with respect to the number and quality of labels that

Table 1. Summary of co-branding literature.

Main authors	Types of brand/label	Key observations
Geylani et al. (2008); Park et al. (1996)	Individual brand (co-branding)	Increase in value if complementarity between attributes or associated brand images. Threshold effect if high perceived value of the brand.
Barone et al. (2000); Gupta and Pirsch (2006)	Individual brand and non-profit association brand (cause marketing)	Increased value if perceived consistency of the brand association; otherwise counter-productive.
Bauer et al. (2012); Hassan and Monier (2006); Larceneux et al. (2012)	Organic label and individual brand Public quality labels and individual brand	The premium for the label depends on the strength of the individual brand image.
Dekhili and Achabou (2013); Janssen and Hamm (2012); Sirieix et al. (2013); Tagbata and Sirieix (2010)	2 labels	Increased value, in terms of preference or choice, for dual labelling when the labels are perceived as complementary or consistent. Does not always lead to a premium effect.

could be combined, and their potential price impact. Alter Eco, for example, recently decided to return to greater packaging simplicity, while the company Saveurs et Nature, fearing information overload, sometimes chooses not to publish all of its awarded certification logos on the packaging of its organic chocolates.

Given this context, the originality of the present research is to conduct an experiment using three labels, all of which enjoy excellent reputations and a strong level of consumer confidence. The selected labels cover three different aspects of the product: its organoleptic qualities (Label Rouge, LR), its organic, ecological focus (Agriculture Biologique, AB), and its eco-social policy (Fairtrade/Max Havelaar, MH). We will study the increase in value for the consumer (measured in terms of WTP) of a multi-labelling strategy including up to three labels on the packaging of a specific food product. The evolution of this perceived increased value of the product will then be examined in relation to the different combinations of labels and the number of labels on the packaging.

Our research is structured and developed in five parts. In the first part, we will present a review of the literature on the monetary valuation of labelled products by consumers, and the effects of the interactions between brands and labels on consumer's WTP. The second part will explain the study context and the hypotheses to be tested. The third part will describe the methodology used and specify the practical combinations of the three labels studied within the context of a specific food product: honey. The fourth part will discuss the findings, highlighting the theoretical contributions of the study and suggesting future avenues for research. Finally, the fifth part will address the managerial implications of the study and the limitations of the research.

Effects of multi-labelling on consumer behaviour

The existence of possible complementarities and substitutabilities perceived by consumers in the presence of several brands and logos displayed simultaneously on a given packaging has been studied in various fields of marketing. The field of co-branding has particularly focused on the circumstances in which the combining of two brands can increase or decrease value. Cause-related marketing (the cooperative efforts of a for-profit business and a non-profit organization) has made it possible to study the perceived coherence effect of brand associations (belonging to different organizations) on consumer purchasing behaviour (Table 1).

More directly related to the issue at hand, several recent studies have analysed the effects of combining brands, logos or labels (complementary or not), by specifically examining the different product characteristics. Bernard and Bernard (2009) measured willingness to pay for organic milk in the USA (154 US respondents using a method inspired by the Vickrey auction[4]) and compared it to the WTP measured when only two criteria of the organic label were communicated to the consumers. They showed that the premium attributed in this case was not significantly different from that attributed to the product after all of the criteria of the specifications for the organic label concerned had been made known. They concluded that additional characteristics resulted in a diminishing marginal utility for added attributes on milk: consumers were willing to pay less for groups of characteristics than for each characteristic individually. Larceneux et al. (2012) also contributed to the understanding of the marginal utility of additional characteristics in the case of food products with an organic label and brand logo on their packaging. In an experimental study conducted with 122 consumers present in an actual store, the authors showed that an increased WTP related to the presence of an organic label on the packaging was significantly greater for brands that were not highly valued by the consumers. The higher the brand equity, the lower the premium associated with the presence of the organic label, even to the extent of being zero for a brand with a very strong image. In an empirical study carried out in Germany, Bauer et al. (2012) also showed that the premium related to the presence of the German organic label varied depending on the brand of the product on which the label was affixed. This premium was greater for a retailer's brand than for other brands. Hassan and Monier-Dilhan (2006) found similar results in their analysis of a database on the consumption of milk, yoghurt, eggs, Camembert and ham by 8000 French consumers. They determined that the more the associated brand enjoyed a high brand equity, the lower the premium attributed for the presence of several public quality labels (PGI, PDO (Protected Designation of Origin), AB and LR). De Pelsmacker et al. (2005b) reported a lack of such impact for ethical labels.

Finally, certain studies have examined the situation where quality signs on packaging comprise two labels. Janssen and Hamm (2012) compared consumer WTP in six European countries for competing organic labels affixed on the packaging of apples and eggs in addition to the official European Union (EU) organic label. They showed that consumers often preferred products carrying a logo explicitly containing the term 'organic' in addition to the generic logo corresponding to the European label. WTP was also higher. And although the two labels guaranteed equivalent quality, their dual presence tended to increase product value in the minds of the consumers. Dekhili and Achabou (2013) confirmed the potential increase in value of dual labelling by showing, on the basis of a conjoint analysis, that the addition of a (fictional) eco label on eggs was seen as complementary to the AB label. The combination of both labels thus resulted in a higher consumer valuation of the product, the eco label providing an environmental dimension to complement the AB label, better known for its 'health' and 'quality' dimensions, even though it does include an environmental component. This study did not, however, conclude that the increased value for dual labelling was greater than the sum of that for each label individually. These findings were partially supported by Tagbata and Sirieix (2010). In an experiment conducted on 102 consumers (BDM lottery), the authors showed that the dual labelling of a bar of chocolate ('organic' and 'fair trade') increased the value of the chocolate for the consumers, but resulted in a lower willingness to pay than the sum of the WTP for each of the labels individually. Finally, Sirieix et al. (2013), in a qualitative study carried out in the UK, examined the perceptions of consumers when they were presented with combinations of pairs of sustainable labels (13 labels related to environmental, social, nutritional, biological and origin aspects). The authors showed that dual labelling was appreciated when the labels' characteristics were perceived to be complementary. However, certain combinations of labels could be rejected and actually lower value, either due to the negative impact of one under-valued label on the other, or to perceived contradictions between the labels in terms of credibility or message.

In summary, three fields of observation can be used as data sources to assess the impact of multi-labelling on consumer WTP (Table 1):

- *Research analysing the effects of the co-branding or not of trademarks*: These studies emphasize the need for complementarity and coherence between the two brands.
- *Research analysing the effects of the combination of an individual brand and a label*: Studies of this nature highlight the fact that the premium attributed to the label depends on both consumer knowledge of this label (knowledge based on a small number of criteria) and the relatively weak image of the brand with which it is associated.
- *Finally, research, still very uncommon today, analysing the effect of dual labelling, without reference to a specific individual brand*: In this case, the impact would tend to be positive in terms of choice and perceptions when the label combination is perceived to be coherent. There are still questions surrounding the premium effect of dual labelling, and it is mentioned in only a few studies with partly contradictory findings.

The present study proposes to analyse the conditions of this premium effect through the progressive increase in the number of logos on packaging, focusing on three well-known labels and without reference to a specific individual brand. The failure to examine more than consumer willingness to pay is undoubtedly a limitation, but this can be justified by the importance of this factor in company labelling policies. Indeed, the additional criteria related to new labelling generate additional costs. Given the highly competitive nature of the agri-food industry, producers can only adopt these labels if the ensuing increase in value for consumers is greater than the additional costs involved. Moreover, the price issue is all the more essential inasmuch as it interacts with consumer confidence: as Hamzaoui-Essoussi et al. (2013) showed, better pricing, corresponding to the real added value of a label, is a factor promoting trust. The right price is therefore a key issue.

Study context and research hypotheses

This research aims to study the evolution of the premium attributed according to the number of labels used on the packaging of a specific food product: honey.

Label choice and food product used for the empirical study

To examine the impact of multi-labelling on consumer WTP, the aforementioned studies encourage a focus on both choice of label and food product used.

Concerning label choice, previous studies highlight the difficulty of implementing co-branding strategies (individual brand and/or certification mark) when the logos and labels concerned appear asymmetrical in terms of their brand image, reputation or trust. Label combination presents certain significant risks, notably that the presence of one of the labels will decrease the value of the other. Therefore, this study sets out to go beyond these observations to examine the evolution in the premium attributed for different label combinations, all of which send out strong value signals, and enjoy similar reputations and degrees of trust. Three food product labels were selected for this study: Fairtrade/Max Havelaar (MH), Agriculture Biologique (AB), and Label Rouge (LR).

The three main reasons for choosing these labels

First, all three labels have a solid reputation and enjoy a very high level of consumer confidence in France.[5] The relative homogeneity in the strength of the quality signal coexisting on the packaging allows focus to be placed on the label's underlying message. The effects of complementarity or substitutability between these labels can be at least in part explained by knowledge of the information contained in their specifications and the meaning associated with each of them.

Moreover, these three labels have specifications geared towards different and potentially complementary aspects (Sirieix et al., 2013). It can

therefore be assumed that they provide consumers with assurances of different values. Drawing on the work of Aurier et al. (2004), Gabriel and Urien (2006), Holbrook (1999), and Rivière and Mencarelli (2012), the different dimensions making up the perceived value of an offer can be grouped under four functions: (i) instrumental, driven by the utilitarian or economic dimension of the value; (ii) identity, highlighted by the self-expression or social dimension of the value; (iii) hedonic, including the experiential or aesthetic dimension of the value; and finally, (iv) ego-political, represented by the ethical or spiritual dimension, by integrity or more generally the responsibility or accountability conveyed by the consumption value. The LR label's value promise is more instrumental and self-oriented (emphasis on aspects such as safety, quality, and manufacturing method). The MH label adopts an identity (expression of social ties, oriented towards others), or ego-political function, putting forward the ethical dimension of a social practice. The value promise associated with the AB label appears more multifaceted compared with the other two (Dekhili and Achabou, 2013). It covers both expression and ego-political functions, as well as a hedonic function given that many consumers believe AB products to be more tasty.[6]

Finally, while these labels have different specificities, they nevertheless have certain points in common. Generally, the MH and AB labels are both oriented towards environmental protection. Moreover, beyond the economic and social aspects that represent the uniqueness of the MH label, its stated intention is to encourage small producers in the South to adopt criteria for organic production. LR also shares principles with the other two specifications concerning production practices (even though it is more oriented towards the organoleptic quality of its products). In the case of honey, similarities between the three labels indeed exist, mainly with regard to contamination restrictions, water content and hydroxymethylfurfural content,[7] although the respective authorized levels vary significantly. Thus, the addition of these labels results in the emergence not only of redundancies likely to generate a sub-additive of valuations, but also the risk of conflicting messages in the minds of the consumers.

To analyse the effects between labels with the least possible bias, the studies outlined above recommend that an individual brand should not be specified. Nevertheless, product appearance must be coherent with the labels affixed on the packaging. Hence our decision to use honey, a product consumed by a large majority of the French population,[8] and the packaging of which regularly includes various logos. Indeed, there are many possible manufacturing processes for this product, which largely determine its organoleptic qualities as well as its effects on the environment and on human health. Within this context, one characteristic of honey is that the specifications of the three labels studied are heavily focused on post-harvest transformation processes.[9] In addition, in the protocols where it was presented, the jar gave the appearance of a good quality honey, but without any particular distinction: a standard, neutral-coloured, 500 g glass jar (this is one of the quality criteria mentioned by consumers in surveys), with the indication 'artisanal blossom honey' on the label.

Research hypotheses

One specificity of the proposed study was to observe the progressive impact on consumer WTP of the gradual addition of labels (each enjoying a similar reputation and level of trust), up to a total of three labels on the packaging. Because these labels have different specifications, they were all likely to provide additional value to the consumer, resulting in a premium for the product.

> Hypothesis 1: The added-value of multi-labelling: the premium for an additional label on a food product is always strictly positive (due to the additional characteristics communicated).

Each of the labels chosen within the study framework offers different value dimensions for the consumer. These dimensions refer to potentially complementary functions, thereby resulting in overall increased value from combinations of these labels. This positive effect of co-branding is highlighted in the majority of the studies cited above. Most of them only observe this effect, however, when the packaging changes from a product logo to

a product logo plus a label. It would be interesting to validate the stability of this finding in different packaging examples, including a maximum of three labels, which is becoming increasingly common for certain food products.

The product's increased value due to multi-labelling may be tempered, however, by two factors: first, the information contained in the labels and their specifications, part of which is redundant; and, second, the cognitive abilities of consumers who are hindered by an increase in the amount of information provided by the accumulated labels. These two factors are evaluated under two additional hypotheses.

> Hypothesis 2: Substitutability and redundancy effects: the premium attributed for a combination of labels is less than the sum of the premiums for each of the labels individually.

Combined with our first hypothesis, this means that the premium for a pair of labels will be greater than the premium for each of these labels individually, but less than the sum of the two. The same is true in the case of triple labelling: the premium obtained in this case will be greater than the premium for dual labelling, but less than the sum of the premiums for each individual label (and also less than the premium for dual labelling plus that for the remaining individual label).

This hypothesis can be justified mainly by the substitutability and redundancy between specifications, which are the opposite of complementarity effects. As highlighted in the literature review, the findings on this subject are not stable. Some previous research is consistent with the hypothesis presented (Bernard and Bernard, 2009; Tagbata and Sirieix, 2010, although in both cases the research was limited to two labels). Conversely, Dekhili and Achabou (2013) perceived a mutually reinforcing effect between an environmental label and an organic label.

> Hypothesis 3: Marginal decrease in value and information overload: the additional premium becomes lower with the greater the number of labels used on the food label.

This third hypothesis, which to our knowledge remains untested to date, observes the quantitative impact of information on the additional premium, thus completing the study of an effect that could, in contrast, be described as qualitative in our second hypothesis.

It can be justified by the demonstration of a threshold effect for co-branding (Park et al., 1996), and by a number of studies on the adverse effects of an excessive amount of information provided to consumers. Bourgherara, Grolleau and Mzoughi (2007) showed that packagings that were overloaded with information on the products' environmental quality had counter-productive effects on consumers, who were no longer willing or able to read and process this information. Within this framework, the authors demonstrated the relevance of the Yerkes-Dodson law in psychology, which assumes the existence of a maximum level of information beyond which additional information has a detrimental effect on the quality of the consumer's decision-making process rather than improving it, because of the complexity of the cognitive processing involved. In a similar vein, Bernard and Bernard (2009) showed that consumers attributed a decreasing value to the additional characteristics communicated by the organic label on milk. It could, then, be reasonable to assume that supplementary signs of quality provided by displaying another label, especially in third position, would encourage the consumer to simplify the information, taking only some elements seen as representative of each of the labels, and thereby facilitating the overall reading of an increasingly complex food packaging label.

Research Methodology

The hypotheses were tested using data acquired from a quantitative empirical study to measure WTP for the labels and their combinations.

Choice of methods for measuring WTP

Willingness to pay, defined as the maximum price a consumer is willing to pay for a product or service, can be estimated using either survey data, notably from contingent valuation (Mitchell and Carson, 1989) or conjoint analysis (Green and Srinivasan, 1990), or bid response data for an incentive,

second-price auction (Vickrey, 1961) or a Becker, DeGroot and Marschak (BDM) lottery (Becker et al., 1964) (for an overview of these methods, see Le Gall-Ely, 2009).

Two methods have been used in this study: contingent valuation and BDM lottery. While it is easy to use, the drawback of the contingent valuation method is that it offers little incentive for respondents to reveal their true WTP (Völckner, 2006; Wertenbrock and Skiera, 2002). Without actual purchase of the product, it suffers from hypothetical bias. A strategic bias may also occur when respondents deliberately shape their responses in an attempt to influence the survey results to suit their own interests. They do this either by overestimating their WTP (to affect the placing on the market of the good or service, to please the interviewer, or to avoid expressing their preference for a lower quality product at a lower price), or by underestimating their WTP (in the hope that a lower sales price will be established for the product). However, the payment card approach enables some of these biases to be limited (Mitchell and Carson, 1984).

In the BDM lottery, the participating respondent establishes the maximum price at which they would buy the proposed product. The selling price is then determined randomly (e.g. by drawing a price-marked ball from a box). If the price drawn is less than or equal to the expressed WTP, the participant purchases the product at this drawn price. Otherwise, the participant cannot purchase the product (Becker et al., 1964). The respondent is thus placed in a situation where the amount expressed will not affect the sales price. Theoretically, it is in the interest of a rational participant to reveal their true WTP, thereby limiting the emergence of strategic bias (Kagel, 1995; McAfee and McMillan, 1987; Shogren et al., 2001). However, this method is somewhat removed from the decision-making process implemented by a buyer in a store (Hoffman et al., 1993). In research comparing panel data (an international study concerning over 20,000 products) and declarative data, Miller et al. (2011) recently showed, however, that the findings obtained by the BDM method were the closest to actual observed consumer behaviour, and that even though the findings from contingent valuation deviated from this behaviour, the difference did not ultimately result in different marketing decisions.

However, no method currently exists for measuring WTP in a way that ensures the validity and reliability of the findings (Le Gall-Ely, 2009). By using two methods for measuring WTP (contingent valuation and the BDM lottery), our empirical research proposed to study consumer WTP for these different combinations of labels. The aim was not to obtain a measurement of WTP per se, nor to compare the findings from the different methods, but rather to get as close as possible to the WTP differentials for the various combinations of labels.

Implementation of the three studies

The three studies were carried out in France in 2012. The first two (online survey and face-to-face interviews) used contingent valuation to measure WTP for different offers of labelled honey (AB, LR, MH, AB+MH, AB+LR, MH+LR, or AB+MH+LR). The main difference between these two studies lay in the degree to which reference price information for the honey was provided to the respondents before they answered the evaluation questions. In the first study (contingent valuation with 307 respondents), the average market price for a quality honey (€5), established after in-store verification, was communicated. The respondents were then asked to note, on a payment card (€5, €5.50, €6, €6.50, €7, €7.50, €8 and more), the price at which they would be willing to purchase the honey, depending on the different combinations of labels. In the second study (contingent valuation with 108 respondents), no indication of the initial price was given. The third study (BDM lottery with 103 respondents), conducted via face-to-face interviews, also measured WTP with no prior indication of the price information. Moreover, the last two studies included slightly fewer questions. They were intended to be compared with the central findings of the first study (concerning WTP) conducted a few weeks earlier.

Presentation of the results

Profile of the honey consumers surveyed

The profile of the honey buyers who participated in the survey was relatively heterogeneous for each of

Table 2. Consumer WTP for the different labels and their combinations.

	Contingent valuation 1 Given ref. price of €5 *n* = 307 Average/Standarddeviation	Contingent valuation 2 No reference price *n* = 108	BDM lottery No reference price *n* = 103
AB	6.03 / 0.79	5.54 / 1.30	5.70 / 2.15
MH	5.83 / 0.73	5.32 / 1.36	5.59 / 1.85
LR	5.67 / 0.68	5.28 / 1.16	5.60 / 1.74
AB–MH	6.15 / 0.81	5.89 / 1.40	6.32 / 2.15
AB–LR	6.06 / 0.80	5.93 / 1.29	6.30 / 2.09
MH–LR	5.92 / 0.74	5.78 / 1.36	6.33 / 1.75
AB–MH–LR	6.27 / 0.86	6.37 / 1.59	6.79 / 2.13

the samples (Appendix 2), mainly in terms of levels of education and socio-professional categories. These were much higher in the first study than in the other two because of the ~~manner in which~~ the respondents were approached (voluntary response for the CV1 study, face-to-face answers in a public place for the CV2 and BDM lottery studies).

However, consumer perceptions of the labels were quite similar, which is consistent with several other studies showing that the perception of labels is not very dependent on socio-demographic characteristics (Organisation for Economic Co-operation and Development (OECD), 2011). For these buyers, quality was first defined by the origin of the product (roughly two-thirds of the respondents), then its label (a little more than half), and finally its means of production (about half). The labels were generally perceived as quite credible, offering a guarantee of quality and trust for the majority of respondents. Indeed, only 25% of the consumers surveyed considered the labels to be nothing more than a marketing strategy, and only 2% considered that they served no purpose whatsoever. Questioned directly on the issue, nearly three-quarters of the respondents said they trusted the AB label, and more than two-thirds trusted the LR and MH labels. The AB label was also the first choice at equivalent prices (Appendix 2). The majority of respondents also believed that the three labels were strictly regulated and controlled. Finally, they also had a fair amount of knowledge about the three labels, with just under two-thirds of them able to correctly identify the main characteristics of the specifications. All of these findings are similar to those of national surveys on the subject (IFOP, 2010; IPSOS, 2012; OECD, 2011; Tavoularis et al., 2007 for the CRÉDOC).

It is important to note that the reputation of and trust in all three labels were high (with a slight advantage for the AB label). This symmetry was taken into account in the analysis of WTP for these three labels and their combinations.

WTP for the different labels

Table 2 presents the average WTP for the different label combinations. The findings highlight variations in terms of levels of WTP depending on the methods used (differences that are sometimes significant according to Tukey's test (Appendix 3), but with comparable orders of magnitude). In line with the findings from previous research, the majority of respondents expressed a positive WTP for each of the three labels presented individually, the average increase in value for a label ranging from 13% to 20% when a reference price for conventional honey was provided (between 16% and 19% for 84% of respondents). The AB label systematically generated the highest average WTP, while the LR label tended to have the lowest average WTP.

Table 3 specifically highlights the premium after adding labels, depending on the label that has been affixed first. Overall, it appears that the premium for each combination of labels was positive. Figure 1 enables a good visualization of the evolution of average WTP according to the presence of one, two or three labels, depending on the three measurements. While the curves appear fairly similar, it can

Table 3. Average marginal premium for label additions.

	CV 1	CV 2	BDM		CV 1	CV 2	BDM		CV 1	CV 2	BDM
Label: AB	1.03			**Label: MH**	0.83			**Label: LR**	0.67		
Addition 2nd label*	0.08	0.37	**0.61**	**Addition 2nd label**	0.21	0.52	**0.73**	**Addition 2nd label**	0.32	0.58	**0.72**
Addition 3rd label	0.16	0.46	**0.48**	**Addition 3rd label**	0.23	0.53	**0.47**	**Addition 3rd label**	0.28	0.51	**0.47**

*When taking two labels into account, we have used average consumer WTP for both proposals (e.g. AB and MH + AB and LR).

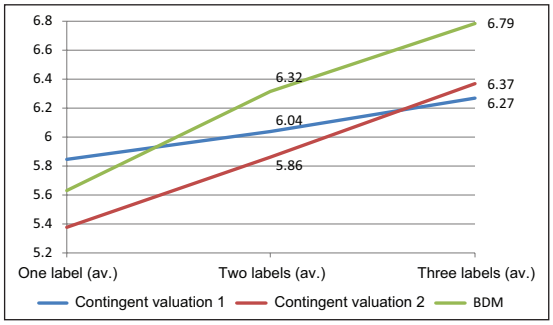

Figure 1. Average evolution of the premiums according to measurement method.

be seen that WTP evolution was relatively linear with both contingent valuations (CV), while the marginal increase in value decreased in the case of the BDM lottery.

Validation of the hypotheses according to measurement method

The research hypotheses cannot be validated in the same way depending on the methods used.

Contingent valuation 1 (CV1), for 307 respondents, **partially confirms Hypothesis 1**. Taking into consideration the average WTP for one and two labels respectively, the premium was certainly significant for the transition from one to two labels, and similarly for the transition from two to three labels (Hypothesis 1 validated, the mean test of the paired samples are significant [$p < 0.000$]). But when each configuration is studied in detail (Table 1), certain label additions did not result in a premium. This was the case for the transition from AB to AB–MH, and from MH to MH–LR. The transition from AB–MH to AB–MH–LR may be considered as generating a premium but only at a risk level of 5.7%. The other eight configurations of label additions resulted in a significant premium (threshold of 1%). **Hypothesis 2, on the other hand, is validated for every label combination**. The average premium attributed in the case of combined labels (no matter what they were) was significantly lower than the sum of the average premiums for each of the same labels individually (Table 2). While the average premium attributed for one label was €0.84, it was on average €1.05 for two labels and €1.27 for three labels. It should be noted, however, that the combination MH+LR, the association of the two least valued labels individually, generated the strongest increase in value. On the contrary, the lowest additional gain was found when combining the AB label with one of the other labels. Finally, it would seem that the addition of a third label was not less valued than the addition of a second: **Hypothesis 3 therefore cannot be validated**. Indeed, Table 2 shows that, for CV1, the increase in value for a second label varied from €0.08 to €0.32 depending on the combinations, and from €0.16 to €0.28 for a third label. The difference between the average premiums associated with the transition from one to two labels and from two to three labels is not significant (two-way ANOVA).

Contingent valuation 2 (CV2) was based on the same methodology as CV1 (except that the questionnaire was completed face-to-face) and was used mainly to check whether the findings from CV1 were the same when initial price was not provided. Logically, variance should be greater in the absence of a common reference.

But the prices announced by the respondents were not considerably different from those obtained when a base price of €5 was announced (CV1). First, the WTP levels obtained by this survey differed from those of CV1 for single-label configurations (again, however, the AB label had the highest value, followed by MH and finally LR), but not in the case of multi-labelling (Tukey's test, Appendix 3). The increase in value for the different two- or three-label configurations was the same for these first two surveys. In addition, the findings were similar for the three hypotheses (Table 2): **Hypothesis 1 is confirmed here** – the addition of labels on the jar of honey systematically resulted in the attribution of a premium. **Hypothesis 2 is also confirmed**. **Finally, CV2, like CV1, results in a rejection of Hypothesis 3**.

Compared with the first two studies, the findings are slightly different for the *BDM lottery* from a cardinal point of view: dual and triple labelling sometimes resulted in a significant increase in value (Appendix 3) in this case. But the level of the declared prices remained around €6 (which confirms that the respondents had a good understanding of the value of the honey). From an ordinal point of view, the BDM lottery also resulted in the unequivocal validation of **Hypotheses 1 and 2**. However, **Hypothesis 3 is also confirmed here**, contrary to what was observed in the contingent valuation studies. The affixing of a third label tended to generate a premium that was significantly lower (€0.47 on average, Table 3) than that attributed for a second label (€0.69 on average) (two-way ANOVA, threshold of only 5%). The BDM lottery, which was expected, in principle, to generate a more realistic measurement of WTP (Miller et al., 2011) thus contradicts the findings of the contingent valuation surveys. **It is therefore impossible to reject Hypothesis 3**.

Discussion of the results and theoretical contributions

The results obtained were analysed and discussed in light of the hypotheses. Table 4 summarizes all of the theoretical contributions and possible conceptual extensions.

Discussion of Hypothesis 1: The role of trust, mistrust, feelings and emotions

It would appear that, on average (and also in each individual case with only two exceptions in CV1), consumers were willing to pay a significant premium for each additional label on the jar of honey, up to a total of three labels. Whatever the method used, the average WTP for three labels was greater than the average WTP for two labels, which was itself higher than the average WTP for only one label. Hypothesis 1 is therefore validated by all three studies.

More specifically, there are certain disparities within our samples, marking specific consumer segments. A breakdown of the respondents by the way in which they attributed more value depending on the combination of labels was established. In particular, we identified that 36% of respondents did not attribute more value for the combination of labels (rejection of Hypothesis 1). Among this group, certain respondents considered that none of the labels presented increased the value of the honey (around 15%). Others considered that one label increased the value of the honey, but that associating it with one or two other labels did nothing to further increase value (less than 20%). Finally, some respondents (3%) even considered that the combination of labels took value away from the product. The remaining 64% of respondents generally considered that the labels and their combinations contributed to increase the value of the honey.

The reason for this does not seem related to the additional information on the specific characteristics provided by each label (conditions of production, ingredients used or expected taste). Indeed, the WTP for the labels/combinations of labels cannot be explained in a statistically significant way by the sociocultural profile of the respondents, their level of knowledge of the labels, their consumption habits, or their consumption selection criteria. The level of the respondents' actual knowledge was measured using a list of 21 questions concerning the content of each label. The findings showed that those who were very familiar with the labels were not necessarily more inclined to attribute a premium. There was nevertheless a variable correlated with WTP:

Table 4. Theoretical contributions related to the hypotheses tested and possible conceptual extensions.

Statement of hypotheses	Validation	Findings	*Suggested research avenues*
H1: The premium for an additional label on a food product is always strictly positive (due to the additional characteristics communicated).	Yes (average)	• The accepted premium is not directly related to either the level of awareness of the labels, or the pro-environmental behaviour of the consumers (cognitive and conative factors). • The refusal to attribute a premium is correlated with a lack of trust in the labels. • The AB label is the one for which the respondents attributed the highest premium and displayed the greatest trust.	• *The motives for trust in the labels (credibility, control) and lack of trust (uncertainty, scepticism) result in a different expression in value attributed to a set of labels. Uncertainty generates a cost/benefit-type economic evaluation, while credibility tends to generate a feeling (moral value).* • *Rather than representing the expression of a rational cost/benefit-type economic calculation, the economic evaluation of a set of environmental labels relies to a great extent on feelings; it expresses values of an ego-political or expression function.*
H2: The premium attributed to a combination of labels is less than the sum of the premiums attributed to each of the labels individually.	Yes	• The relative weakness of the premium is related to the level of value of the original label: the most highly valued label (AB) was the one that benefited the least from being associated with one or other of the other two labels. This effect was the opposite for the LR label, the least valued of the three. • The combinations that increased value concerned labels that were somewhat redundant in terms of information.	• *Embedding effects: environmental labels have wider consumer value expression functions than the strict specifications to which they are subject, amplifying the effects of substitution when they are combined.* • *The AB label's fuller promise of value, associated with the higher credibility of this label compared with the others, highlights the prototypicality characteristic of this label.*
H3: The additional premium decreases with the increase in the number of labels used on the food packaging (information threshold effect).	No	• Contingent valuations: the premium attributed to a third label did not differ from that attributed to the second. • BDM: The premium attributed to a third label was high but less than that attributed to the second. • This premium was attributed even by people considering the three labels to be broadly equivalent, and who did not attribute an increase in value for label combinations.	• *For 'convinced' consumers (those who already attribute a premium in the case of dual labelling): reinforcing effect of the perceived congruence by the addition of a third label, different from the other two and of a similar reputation. This congruence leads to a greater overall perceived value (three labels).* • *For consumers initially less convinced: the multiplication of corroborative and credible evidence (reputable labels) is necessary in the case of uncertainty where credence attributes are concerned.*

trust. The choice of *not* attributing a premium can be explained by a general lack of trust in the labels. Indeed, one significant finding is that the respondents who did not recognize an increase in value as a result of the labels were those who admitted they had a lower degree of trust in said labels.[10] But conversely, our findings do not show that those who did trust the labels increased the value of the honey.

These findings raise the following two discussion points (future avenues for research) around this notion of trust.

1) The lack of trust in environmental labels changes when these labels are associated with a food product. Trust is dependent on the label's credibility and underlying control (Sirieix et al., 2013) and can be expressed as the decision to 'just assume that food is safe' (Green et al., 2005: 525). Lack of trust in the food sector is linked to uncertainty, scepticism and risk (Chen, 2011). It would seem, on the one hand, that those consumers who trusted the labels were actually no more likely than the others to consider the credibility of the labels as justifying an additional price when combined. On the other hand, the uncertainty on the part of the more sceptical consumers was not fully resolved either by combining the labels, thereby justifying the lack of premium for this group.

One possible explanation is that, for some consumers, the credibility given to each label was such that it could not be enhanced through combination with other labels. Like some associations between a recognized label and a strong brand, which lead to no premium, or only a moderate one (Bauer et al., 2012; Hassan and Monier-Dilhan, 2006), there was no extra credibility resulting in greater economic value. For those consumers with a low degree of trust in the labels, the combination of the latter amounted to an accumulation of doubts rather than an increase in the number of positive elements to foster confidence. In the cost/benefit calculation of an economic value, the costs (or uncertainty) remain; a situation that does not lead to the attribution of a premium. Sirieix et al. (2013) thus observed that the combination of two opposite labels in terms of credibility or value could result in an overall negative evaluation.

2) The economic increase in value from label combinations is not directly related to either cognitive factors (degree of knowledge of the labels) or conative ones (previous purchasing behaviour or decision making). This finding, counter-intuitive in principle, was observed in another form by Ehrich and Irwin (2005). They discovered that the consumers who were the most sensitive to environmental protection issues did not necessarily take the trouble to find out more about the manufacture of a product in order to avoid a possible negative emotion. This behaviour was all the more notable the greater the perceived attractiveness of the product. This could mean, for our study, that the economic evaluation of a combination of environmental labels was not based on the additional and aggregate information resulting from this association, but rather on emotional factors. This interpretation is in line with Chen (2011) and Chernev and Gal (2010) on the subject of food products. The combination of labels cannot be directly expressed in terms of additional costs or benefits, promoting the possibilities of a rational calculation of value. Instead, it would tend to strengthen feelings and emotions by expressing non-utilitarian values.

Discussion of Hypothesis 2: Embedding effect and prototypicality

While the additional information provided by a new label did, on average, increase product value in the minds of the consumers (Hypothesis 1), the premium attributed was nevertheless moderated by the effects of substitutability between the labels. In other words, the overall value of a combination of labels is not the mere aggregate of the values assigned to each of the labels in the combination. This finding reinforces that of Tagbata and Sirieix (2010), among others.

Table 2 allows us to further extend the analysis of this effect by showing differences according to

label combination. The AB label, the one that was most highly valued individually, was also the one that derived the least benefit from its combination with the other two labels. Conversely, the label that benefitted the most from being combined with a second label was the LR label, the one with the least individual value (which is consistent with the fact that, at the same price, it was the least often purchased). LR+AB or LR+MH were thus valued almost as highly as the AB–MH combination (no significant difference according to CV2 and the BDM lottery). A LR honey would therefore have a lot to gain from the acquisition of a second label with different specifications (more oriented towards the others in this case). All in all, the sub-additivity effect was lower for the labels that were initially less well valued, even though they benefitted from a solid reputation.[11] This is consistent with the findings of Sirieix et al. (2013) who showed that the most popular consumer combinations involved associations between labels that were both well recognized and perceived to be complementary (the MH label, the most popular one in their investigation, was quoted in many combinations).

This finding raises questions on the reasons for the higher substitution effect observed for the AB label. Two discussion points emerge, representing possible interpretations to be explored in future research:

1) Irwin and Scattone-Spira (1997) attributed to an embedding effect the fact that a set of environmental attributes (in their case, recycling and biodegradability) resulted in a lower overall rating than the sum of each taken separately. They noted that if these attributes were not correlated, this effect was an anomaly, substitutabilities being logically absent. For these authors, the explanation for this lay in the fact that environmental offers were evaluated, not on a cost/benefit basis, but on the capacity of the offer to express an ethical attitude. Thus, the presence of one or more environmental labels would tend to stimulate the same ethical implication. The general idea of 'doing one's part' implies that attributes representing the same value (the environment) do not exhibit a combination that is economically more valued.

At least two corroborative elements give credence to this explanation for our research. On the one hand, the discussion of Hypothesis 1 above highlights the lesser role played by cognitive and conative factors in the economic increased value from label combinations. On the other hand, the values associated with the labels in our study are more than just utilitarian. This is particularly the case of the AB label, which, in the minds of the consumers, expresses consumption values that are not strictly related to the label's specifications. Moreover, Hamzaoui-Essoussi et al. (2013) also observed that even for producers of organic products, the objectives were notably related to local economic development and the construction of a social vision of agricultural production; characteristics that are not found in the specifications.

2) The AB label's promise of value extends beyond its specifications. It is the most comprehensive label, covering some of the commitments of the other labels with which it is combined, resulting in a significant substitution effect. The AB label effectively regroups different values of consumption that have both instrumental and expression functions, as well as a hedonic function linked to the expected taste enjoyment. In addition, AB is the label that triggered the highest premium and the greatest degree of trust among respondents (measured using a Likert scale). These different features result in the AB label taking on a distinctive characteristic of prototypicality, serving as a marker for consumers and representative of a broad (given the associated values) 'environment' category. As such, its power of meaning is both larger and wider than the other labels; association with one or other of the latter was therefore logically less valued. Irwin and Scattone (1997) underlined that people who have higher ethical beliefs are more emotionally involved in the embedding effect, and are more likely to exhibit prototypicality judgements.

Discussion of Hypothesis 3: Congruency effect and threshold of persuasion for credence attributes

On average, the additional information provided by a new label increased the product value for the consumer (Hypothesis 1). However, contrary to Hypothesis 3, this premium did not necessarily decrease with an increase in the number of labels on the packaging. The premium for additional labels only decreased for the transition from two to three labels, as opposed to the transition from one to two labels, with the BDM method (and again, with a risk of 4.7%). In this latter case, it nevertheless remained at a significant level of nearly €0.50 (almost 10%). Conversely, in the case of the two contingent valuation studies, the premium attributed for a third label (not greater than for the BDM lottery) was not less than that attributed for a second label. While the BDM method allows for a much better estimate of true WTP than contingent valuation (see Research Methodology above), both of the latter studies (involving more respondents) invalidated this third hypothesis.

It is therefore impossible to confirm Hypothesis 3. However, as in Miller et al. (2011), the difference in outcome between the two contingent valuation studies and the BDM lottery does not result in fundamentally different managerial recommendations. What is clear is that in all cases, the addition of a third label was not neutral and could very favourably influence WTP. This finding leads us to conclude that triple labelling does not reveal a strong threshold effect, as described by Park, Jun and Shocker (1996).

Analysis of the findings reveals two quantitative explanations for this substantial increase in value as a result of adding a third label: an increase in the WTP of some respondents among those who already attributed a price premium for dual labelling; and an increase in the number of respondents attributing a premium with the transition from two to three labels. Each explanation is subject to debate and addresses possible avenues for future research.

1) The increase in value perceived by consumers already attributing a premium for dual labelling may highlight a reinforcing effect. Far from a threshold effect, it would seem these individuals welcomed and accumulated additional information from a third label. It can be assumed that this complementary information gradually offset the effects of substitution observed in Hypothesis 2. This reasoning is supported by the fact that more than 80% of the consumers considered that the labels provided different types of guarantees. Nevertheless, the validation of Hypothesis 2 also highlights the existence of substitutability effects, based on emotional and moral rather than cognitive factors. The following explanation can then be considered: for labels with a strong reputation, complementary and substitutable information could be combined in favour of a perceived congruence, reinforcing each other with the addition of different labels of similar reputation. Perceived congruence is defined as the similarity between two entities that mutually provide meaning to each other, highlighting the relevant and coherent nature of the combination (Maille and Fleck, 2011). This similarity, based on elements (information, organizations) deemed to be credible by the consumers (Kim and Choi, 2012), may stem from the characteristics of each of the labels (their specifications). It may also and more subjectively come from the perceived value (which is consistent with Hypothesis 2) or even similar targets or created social groups (Fleck and Maille, 2010). In future research, it would be interesting to validate the hypothesis of a congruency effect, and to study the origin of the perceived congruence between labels.
2) The second explanation for the observed results is the increase in the number of consumers attributing a premium for three labels. This attracts ‘new’ individuals (for our study, respondents not yet convinced of the additional value of two labels), and thus represents an additional factor of persuasion. In other words, the multiplication of quality signs would tend to have the opposite effect to that of the threshold effect for consumers who do not highly value or have difficulty

distinguishing the differences between these signs. Analysis of the findings shows that a significantly greater number of the respondents who did not attribute an increased value to label combinations reported that they considered the AB, MH and LR labels to be broadly similar (Pearson's chi-squared test, Fisher's exact test). Everything seems to indicate that quality signs had to reach a sufficiently high level (triple labelling) to persuade the person who was uncertain about their value. There are two possible explanations for this. The theory of selective information processing posits that people focus first on the elements of information that they perceive to be consistent with their initial assumption or belief; when corroborative evidence reaches a minimum threshold allowing them to confirm their initial assumption, information processing ceases (Sanbonmatsu et al., 1998). In this respect, the labels, all of a similar reputation, constituted sources of corroborative information, since they were considered equivalent. But on the other hand, these sources of information are uncertain: the labels added to the packaging represented so-called credence attributes (which cannot be verified either before or after product consumption). Decision making based on such attributes requires trust, intuition and emotion (Chen, 2011). However, in cases of uncertainty, more corroborative information is necessary, while the credibility of the source of this information promotes persuasion (Jain and Posovac, 2001). In other words, for those doubting the utility of labels or their combination, the multiplication of labels of a similar reputation and meaning would enable a rather emotional threshold of persuasion (e.g. confidence building) to be crossed and would result in a higher WTP.

Conclusion and managerial implications

This research conducted with three samples of honey buyers using two methods for measuring WTP (contingent valuation and BDM lottery) yielded results that were all the more impressive because they did not depend on the respondents' socio-demographic characteristics, nor on their consumption and purchasing habits, and were not related to a specific individual brand. This enabled several managerial issues to be raised concerning the monetary increase in value of affixing multiple labels on the packaging of a food product. In particular, it should be noted that:

- The marginal premium did not necessarily decrease. Synergies are therefore potentially strong for combinations of credible and recognized labels.
- The LR label was the one that benefitted most from being associated with the others, while the AB label was the one that benefitted the least.
- In the presence of credence attributes, trust with regard to the labels, the level of reputation of the combined labels, and the particularly representative character of one label with respect to the others (resulting in a typicality judgement), are all elements to which producers and companies should pay attention.

From an operational perspective, the main implication of this research is aimed at producers who are already positioned on specific market segments (such as organic or fair trade, for example), and who question whether the costs inherent in obtaining additional certification would be offset by an increase in consumer WTP. By showing that the marginal utility related to the addition of labels does not necessarily decrease, this study recommends the multiplication of labels on packaging when the marginal cost of acquiring additional labels is decreasing. This is the case for honey, given the overlap between the specifications of the three labels considered, but also for other products for which the labels share common constraints (eggs, chocolate, coffee, meat, biscuits, etc.).

As for the actual publishers of labels, who often wonder whether or not to introduce more features behind a given label (for example, social characteristics for the organic label, or organic characteristics

for the LR label), our results advocate targeted, simple labels which can be cumulated.

An additional implication is related to the replacement of the AB label by the European logo, which has a lower level of awareness and trust than that enjoyed by the French label, and is perceived as less complete (OECD, 2011; Zépéda et al., 2013). In the light of our findings, it may be wise for producers to associate it with other complementary labels. This is even more important given the emergence of many other public and private environmental labels with which the EU organic label can be associated. Label combinations and their impact remains a vast field of research in marketing.

Finally, it is important to underline one final contribution of our findings: while each method (contingent valuation and BDM lottery) served to highlight an increase in the average WTP according to label combination, this growth was not of the same order of magnitude according to method. It is worthwhile conducting further research that is specifically directed along the lines of these methodological issues.

The research presented is not without its limitations. First, the empirical studies are exploratory. Empirical research on this issue is still very limited and some caution is called for with respect to the validity of the findings. Additionally, the external validity of the studies undertaken is limited, because our samples of voluntary respondents were not representative of the French population, even though the multiplicity of the methods and the samples allowed a certain number of biases to be eliminated. Moreover, transnational studies show strong national specificities in terms of perception of the labels (Hamzaoui-Essoussi et al., 2013; OECD, 2011), which justifies the scope of this research but limits its external validity. Finally, the findings are not strictly generalizable because only one consumer product was considered (honey). However, analysis of the findings shows that it was the labels themselves that the consumers valued through their WTP rather than the specific attributes of the honey. While the consumers we interviewed on this subject were for the most part very familiar with the general principles of the three labels (roughly two-thirds of the respondents), they knew very little about the honey's specific features (highly technical principles). It would be interesting to generalize the findings obtained by testing the effect of multi-labelling on consumer WTP for other, possibly non-food products.

Acknowledgements

We would like to thank Professors Amina Béji-Bécheur and Nil Özçaglar-Toulouse, as well as our reviewers, for their invaluable suggestions and comments, which helped us improve our article. Our thanks also to Marie-Caroline Hiriart, Jacques Hoornaert, Sophie Larchartre and Nolwenn Roudot for all their help in the collection of data.

Notes

1. See the website: http://www.inpi.fr.
2. The absence of pesticides or genetically modified organisms (GMOs), for example, is a credence attribute (Balineau and Dufeu, 2010) that cannot be identified by mere consumption of the product.
3. See Alter Eco, 2010: 35.
4. In the Vickrey auction (Vickrey, 1961), participants' offers are collected simultaneously and the person who makes the highest bid purchases the product at the second-highest price. This type of auction differs from the traditional English or first-price auction, in which the highest bidder obtains the product at the price they set.
5. According to many surveys. See for example IFOP (French Institute of Public Opinion, 2010), CSA AGENCE BIO (2007) and IPSOS (2012). This is the case for our sample of respondents.
6. An OECD (2011) survey confirmed, in particular, the importance of personal considerations such as health (the first concern for 70% of respondents) in the decision of households to consume these foods. Public policy considerations (environmental protection or animal welfare) play a lesser role. Moreover, the consumers interviewed by Sirieix et al. (2013) often associated an organic label with 'good food'.
7. This is an organic compound derived from the dehydration of the sugars in honey (increased by heat, a low pH-value, and water content), a sign of ageing and poor quality.
8. According to a survey carried out in 2011, honey is consumed by 90% of the French population (APINOV, 2011).
9. The three labels exist for honey, but the combination of the three is theoretical.
10. It should also be noted that we did not find any correlation between the extent of trust assigned to each label separately and the WTP.

11. When AB was combined with MH, the difference in the loss of value with respect to the sum of the separate average premiums was higher (around -€0.71 in the case of CV1 (€1.15 - (€1.03+€0.83)), see Table 2, first column) than when it was combined with LR (-€0.64). There was a more moderate decrease for the MH–LR combination (-€0.42 on average).

References

Alter Eco. (2010) Annual report. Available at: img.altereco.com/site_internet_2012/onglet_altereco/RA_ALTERECO_2010.pdf (accessed 15 June 2014).

Aouina Mejri C and Bhatli D (2014) CSR: Consumer responses to the social quality of private labels. *Journal of Retailing and Consumer Services* 21(3): 357–363.

APINOV. (2011) Enquête de consommation du miel en France. Available at: www.apinov.com/includes/pdf/Enquete_consommation_miel_Apinov_2011.pdf (accessed 26 June 2014).

Aurier P, Evrard Y and N'Goala G (2004) Comprendre et mesurer la valeur du point de vue du consommateur. *Recherche et Applications en marketing* 19(3) : 1–20.

Balineau G and Dufeu I (2010) Are Fair Trade goods credence goods? A new proposal, with French illustrations. *Journal of Business Ethics* 92: 331–345.

Barone MJ, Miyazaki AD and Taylor KA (2000) The influence of cause-related marketing on consumer choice: Does one good turn deserve another? *Journal of the Academy of Marketing Science* 28: 248–262.

Bauer H, Heinrich D and Schäfer D (2012) The effects of organic labels on global, local, and private brands. More hype than substance? *Journal of Business Research* 66(8): 1035–1043.

Becker GM, DeGroot MH and Marschak J (1964) Measuring utility by a single-response sequential method. *Behavioral Science* 9: 226–32.

Bernard JC and Bernard DJ (2009) What is it about organic milk? An experimental analysis. *American Journal of Agricultural Economics* 91(3): 826–836.

Bourgherara D, Grolleau G and Mzoughi N (2007) Is more information always better? An analysis applied to information-based policies for environmental protection. *International Journal of Sustainable Development* 10(3): 197–213.

Chen M-F (2011) Consumer's trust-in-food-safety typology in Taiwan: Food-related lifestyle matters. *Health, Risk & Society* 13(6): 503–526.

Chernev A and Gal D (2010) Categorization effects in value judgments: Averaging bias in evaluating combinations of vices and virtues. *Journal of Marketing Research* 47(4): 738–747.

CSA AGENCE BIO. (2012) Baromètre de consommation et de perception des produits biologiques en France – Etude n°1201517 – Edition 2012. Available at: www.agencebio.org/sites/default/files/upload/documents/4_Chiffres/BarometreConso/Barometre_Agence%20Bio_CSA2012.pdf (accessed 10 December 2013).

De Pelsmacker P, Driesen L and Rayp G (2005a) Do consumers care about ethics? Willingness to pay for fair-trade coffee. *Journal of Consumer Affairs* 39: 363–385.

De Pelsmacker P, Janssens W, Sterckx E and Mielants C (2005b) Consumer preferences for the marketing of ethically labeled coffee. *International Marketing Review* 22: 512–30.

Dekhili S and Achabou MA (2013) Pertinence d'une double labellisation bio/écologique auprès des consommateurs. Une application au cas des œufs. *Economie Rurale* 336: 41–59.

Ehrich K and Irwin J (2005) Willful ignorance in the request for product attribute information. *Journal of Marketing Research* 42(3): 266–277.

Fleck N and Maille V (2010) ~~Trente ans de travaux contradictoires sur l'influence de la congruence perçue par le consommateur: synthèse, limites et voies de recherche. *Recherche et Applications en Marketing* 25(4) : 69–92.~~

Gabriel P and Urien B (2006) Valeurs de consommation et origine territoriale des produits. *Décisions Marketing* 43–44: 41–54.

Geylani T, Inman JJ and Hofstede FT (2008) Image reinforcement or impairment: The effects of cobranding on attribute uncertainty. *Marketing Science* 27: 730–744.

Green PE and Srinivasan V (1990) Conjoint analysis in marketing: New developments with implications for research and practice. *Journal of Marketing* 54(4): 3–19.

Green JM, Draper A, Dowler EA, Fele G, Hagenhoff V, Rusanen M and Rusanen T (2005) Public understanding of food risks in four European countries: A qualitative study. *European Journal of Public Health* 15(5): 523–527.

Gupta S and Pirsch J (2006) The company–cause–customer fit decision in cause-related marketing. *Journal of Consumer Marketing* 23(6): 314–326.

Hamzaoui-Essoussi L, Sirieix L and Zahaf M (2013) Trust orientations in the organic food distribution channels: A comparative study of the Canadian and French markets. *Journal of Retailing and Consumer Services* 20(3): 292–301.

Hassan D and Monier-Dilhan S (2006) National brands and store brands: competition trough public quality labels. *Agribusiness* 22(1): 21–30.

Hoffman E, Menkhaus DJ, Chakravarti D, Field RA and Whipple GD (1993) Using laboratory experimental auctions in marketing research: A case study of new packaging for fresh beef. *Marketing Science* 12(3): 318–338.

Holbrook MB (1999) Introduction to consumer value. In: Holbrook MB (ed.) *Consumer value: A framework for analysis and research*. London & New York, Routledge, pp.1–28.

IFOP (2010) October. Available at: www.rp-net.com/downloader.php?id=151 (accessed 10 December 2013).

IPSOS (2012) July. www.commercequitable.org/images/pdf/donnees_economiques/le_commerce_equitable_en-france_et_dans_le_monde_2012.pdf (accessed 10 December 2013).

Irwin J and Scattone-Spira J (1997) Anomalies in values for consumer goods with environmental attributes. *Journal of Consumer Psychology* 6(4): 339–363.

Jain SP and Posavac S (2001) Prepurchase attribute verifiability, source credibility, and persuasion. *Journal of Consumer Psychology* 11(3): 169–180.

Janssen M and Hamm U (2012) Product labeling in the market for organic food: Consumer preferences and willingness-to-pay for different organic certification logos. *Food Quality and Preference* 25: 9–22.

Kagel JH (1995) Auctions: A survey of experimental research. In: Kagel JH and Roth AE (eds) *The handbook of experimental economics*. Princeton, NJ: Princeton University Press, pp.501–585.

Kim S and Choi SM (2012) An examination of effects of credibility and congruency on consumer responses to banner advertisements. *Journal of Internet Commerce* 11(2): 139–160.

Larceneux F (2001) Critical opinion as a tool in the marketing of cultural products: The experiential label. *International Journal of Arts Management* 3(2): 60–70.

Larceneux F, Benoit-Moreau F and Renaudin V (2012) Why might organic labels fail to influence consumer choices? *Journal of Consumer Policy* 35: 85–104.

Le Gall-Ely M (2009) ~~Définition, mesure et déterminants du consentement à payer du consommateur: synthèse critique et voies de recherche. *Recherche et Applications en Marketing* 24(2) : 91–113.~~

Maille V and Fleck N (2011) ~~Congruence perçue par le consommateur : vers une clarification du concept, de sa formation et de sa mesure. *Recherche et Applications en Marketing* 26(2) : 77–111.~~

McAfee RP and McMillan J (1987) Auctions and bidding. *Journal of Economic Literature* 25: 699–738.

Miller KM, Hofstetter R, Krohmer H and Zhang ZJ (2011) How should consumers' willingness to pay be measured? An empirical comparison of state-of-the-art approaches. *Journal of Marketing Research* 58(1): 172–184.

Mitchell RC and Carson RT (1984) *A contingent valuation estimate of national freshwater benefits: Technical report to the US environmental protection agency*. Washington DC: Resources for the Future.

Mitchell RC and Carson RT (1989) *Using surveys to value public goods: The contingent valuation method*. Washington DC: Resources for the Future.

OECD. (2011) *Politiques de l'environnement et comportement des ménages*. Paris: Éditions OCDE.

Park CW, Jun SY and Shocker AD (1996) Composite branding alliances: An investigation of extension and feedback effects. *Journal of Marketing Research* 23: 453–466.

Rivière A and Mencarelli R (2012) ~~Vers une clarification théorique de la valeur perçue en marketing, *Recherche et Applications en Marketing* 27(3): 97–123.~~

Sanbonmatsu DM, Posavac S, Kardes F and Mantel S (1998) Selective hypothesis testing. *Psychonomic Bulletin and Review* 5(2): 197–220.

Shogren JF, Koo C, Margolis M and List JA (2001) A random nth-price auction. *Journal of Economic Behavior and Organization* 46: 409–421.

Sirieix L, Delanchy M, Remaud H, Zepeda L and Gurviez P (2013) Consumers' perceptions of individual and combined sustainable food labels: A UK pilot investigation. *International Journal of Consumer Studies* 37: 143–151.

Tagbata D and Sirieix L (2010) L'équitable, le bio et le goût. Quels sont les effets de la double labellisation bio- équitable sur le consentement à payer des consommateurs? *Cahiers d'Agriculture* 19(1): ?–?. **[author to provide page numbers]**

Tavoularis G, Recours F and Hebel P (2007) Perception de la qualité et des signes officiels de qualité dans le secteur alimentaire. CREDOC, Cahier de Recherche N° 236.

Vickrey W (1961) Counter speculation, auctions and competitive sealed tenders. *Journal of Finance* 16: 8–37.

Völckner F (2006) An empirical comparison of methods for measuring consumers' willingness to pay. *Marketing Letters* 17(2): 137–149.

Wertenbroch K and Skiera B (2002) Measuring consumer willingness to pay at the point of purchase. *Journal of Marketing Research* 39(2): 228–241.

Zander K and Hamm U (2010) Consumer preferences for additional ethical attributes for organic food. *Food Quality and Preference* 21: 495–503.

Zépéda L, Sirieix L, Pizzaro A, Corderre F and Rodier F (2013) A conceptual framework for analyzing consumers' food label preferences: An exploratory study of sustainability labels in France, Quebec, Spain and the US. *International Journal of Consumer Studies* 37: 605–616.

Appendix 1. Comparison of the characteristics of the LR, AB and MH labels (for honey).

Regulated characteristics	LR	AB	MH
Characteristics			
Content (water, glucose, etc.)/ freshness, T° during storage of the honey, liquefaction; layout of the hives	Strict controls	Strict controls	Incentives to produce according to organic criteria
Allopathic treatments and use of chemicals, feeding of bees	Limited	Prohibited	Limited
Protection of ecosystems and animals		Priority objective	Important objective
Organoleptic qualities		~~Defined and evaluated by a jury~~	
Economic constraints			Minimum small producer purchase price, development subsidy (15%), pre-financing, etc.
Social constraints			Compliance with international labour organization (ILO) rules, partnership, etc.
Measurement and monitoring of these characteristics			
Quantifiability of the attributes concerned in the absence of the label	Research attributes (detectable by the consumption experience)	Credence attributes (undetectable by the consumption experience)	Credence and controversial attributes (uncertain even for experts, Balineau and Dufeu, 2010)
Monitoring principles	Accredited independent body	Accredited independent body	Accredited independent body

Appendix 2. Comparative structure of the three samples of respondents.

		Contingent valuation 1	Contingent valuation 2	BDM lottery
Sex	Proportion of women (%)	69	61	60
Age	Average and range (years)	34 (18–68)	38 (18–76)	37 (18–73)
Marital status	Couple (%)	56	55	52
Children	Proportion having dependent children (%)	46	52	42
Place of residence	Rural vs. urban (%)	33	45	49
Socio-professional category	Employee/mid-level prof.	27	44	41
	Executive	27	6	4
	Student	35	5	7
	Housewife	2	22	20
	Retiree	5	13	10
	Other + not rep.	4	10	18
Educational level	2-year university degree and higher (%)	86	43	45
Purchasing frequency of honey	Less than once a week (%)	72	71	82
	Once a week or more	28	29	18
Jar chosen if identical price	AB (%)	70	60	48
	MH	10	10	19
	LR	9	14	16
	Conventional	10	16	17

Appendix 3: Differences between WTP measured according to method (Tukey's test)

Table A3.1. ANOVA (H0: equal WTP according to method).
(Comprehensive view: do the three methods of measuring WTP result in a globally different assessment?).

Label or combination of labels	Significance
AB	.001
MH	.000
LR	.006
AB–MH	.055
AB–LR	.086
MH–LR	.001
AB–MH–LR	.004

Table A3.2. Multiple comparisons (H0: equal WTP according to method).

Label or combination of labels	WTP assessment method	WTP assessment method	Sign.
AB	CV 2	BDM	.646
		CV 1	.002
	BDM	CV 1	.059
MH	CV 2	BDM	.212
		CV 1	.000
	BDM	CV 1	.171
LR	CV 2	BDM	.085
		CV 1	.004
	BDM	CV 1	.830
AB–MH	CV 2	BDM	.046
		CV 1	.192
	BDM	CV 1	.469
AB–LR	CV 2	BDM	.078
		CV 1	.630
	BDM	CV 1	.198
MH–LR	CV 2	BDM	.002
		CV 1	.527
	BDM	CV 1	.005
AB–MH–LR	CV 2	BDM	.065
		CV 1	.809
	BDM	CV 1	.003