Relative willingness to pay and surplus comparison mechanism

Pierre Combris[1], Eric Giraud-Héraud[2], Alexandra Seabra Pinto[3]

[1] INRA–UR1303 AliSS, Ivry sur Seine, France ; Pierre.Combris@ivry.inra.fr
[2] INRA and GREThA/University of Bordeaux, UMR CNRS 5113, Bordeaux, France: eric.giraud-heraud@u-bordeaux.fr
[3] National Institute for Agrarian and Veterinarian Research (INIAV,I.P.), Oeiras, Portugal: alexandra.pinto@iniav.pt

Paper prepared for presentation at the EAAE-AAEA Joint Seminar
'Consumer Behavior in a Changing World: Food, Culture, Society"

March 25 to 27, 2015
Naples, Italy

Relative willingness to pay and surplus comparison mechanism

Pierre Combris[1], Eric Giraud-Héraud[2], Alexandra Seabra Pinto[3]

143rd Joint EAAE/AAEA Seminar, "Consumer Behavior in a Changing World: Food, Culture and Society", March 25-27, 2015, Naples.

Topic: How to integrate new approaches from social sciences for measuring consumers' behaviour in economics? (objectives 1 and 2 of the 143rd Joint EAAE/AAEA Seminar).

Abstract

We study the relative willingness-to-pay (WTP) of consumers according to the diversity of supply in a market and we show how the presence of substitutes for a given product leads to question the incentive mechanisms commonly used in experimental auctions. We propose a Surplus Comparison Mechanism (SCM) in order to yield WTP estimates which better take into account the choice set available to consumers. After showing the efficiency of this mechanism we test the SCM in a laboratory experiment, reconsidering WTP for food environmental certifications (Integrated Pest Management and Organic certification). It appears that WTPs are decreasing when more alternative certifications are offered to consumers.

Keywords: Experimental Auctions, Willingness to pay, Consumers' surplus, Choice alternatives, Food certification.

JEL classification : C91, D44

[1] INRA–UR1303 AliSS, Ivry sur Seine, France ; Pierre.Combris@ivry.inra.fr
[2]INRA and GREThA/University of Bordeaux, UMR CNRS 5113, Bordeaux, France: eric.giraud-heraud@u-bordeaux.fr
[3]National Institute for Agrarian and Veterinarian Research (INIAV,I.P.), Oeiras, Portugal: alexandra.pinto@iniav.pt

1. Introduction

Very often, the sale of wine may exceed a thousand references in a large supermarket, when a small grocery store can propose only a dozen of brands. Do you think your willingness to pay (WTP) for a bottle of 'Mouton Cadet' would be exactly the same in each outlet? And if you're addicted to port wine, do you really believe that your WTP for the mediocre 2010 vintage would be the same before and after the release of the sublime 2011 vintage? On a local farmers' market of fruits and vegetables, do you think the decision to buy a kilogram of conventional apples may not differ depending on whether or not you have access to the organic version of these apples? And finally, would not a food innovation, such as a GMO product, be more easily accepted by recalcitrant consumers, in the absence of conventional product on the market?

As pointed out by Dan Ariely in his book 'Predictably Irrational' (2008), 'we rely on comparisons' and the decision process of an economic agent is always in the comparison.[1] The presence or absence of substitutes for a product that you wish to purchase can alter the assessment that you make of this product, not only because of the cost of acquiring a substitute when it is not available at the outlet, but also, and perhaps primarily, because of an immediate effect of comparison, often highlighted in marketing research and cognitive psychology. Indeed, the literature on categorization (see for examples Cohen and Basu, 1987; Pothos and Wills. 2011) helps to explain why each individual makes his own selection when faced with a set of several opportunities for making purchases enviable choice. This theory explains how one product can be neglected when it is faced with a large number of alternatives while the same product would be considered in the "category" of enviable products in the absence of this confrontation. Moreover, it is well known in behavioural sciences that the presence or absence of alternatives can influence the choice for a given product. As noticed by Bordalo, Gennaioli and Shleifer (2013, p. 817), the "decoy effect" previously studied by Huber, Payne, and Puto (1982), or the "compromise effect" studied by Simonson (1989)[2] are such examples explaining that, given an offered set, the relative attractiveness of a particular option *x* compared to another *y* often depends on the presence or absence of a third option *z* (see also Bazerman *et al*. (1999) and Tversky and Simonson (1993) for the violation of the *independence of irrelevant alternatives*). Consequently, evidence has

[1] See also the effects of *anchor* and *reference* points, according to the prospect theory (Kahneman and Tversky, 1979).

[2] The decoy effect occurs when the demand for an option is larger when an alternative objectively inferior is added. The compromise effect occurs when the presence of two extreme options favors the choice of an intermediate product.

emerged that decisions are influenced by context, in apparent contradiction with standard assumptions of rational decision making.

We argue that these considerations must lead to revisit the incentive mechanisms used in experimental auctions to elicit willingness-to-pay (WTP) for products. In particular, we think that too little attention has been paid to the impact of the number of goods that participants have to evaluate in an experiment. Fixed costs considerations and the appeal of testing a variety of product attributes justify multiple-good valuation in the same experiment. To avoid demand reduction when participants are in a position to purchase several products, random selection of one product and one treatment is the usual way to proceed, because of this random selection of a binding product, "subjects' bids should not be influenced by the number of goods auctioned" (Lusk *et al.*, 2004, p. 403). However, evaluating the same goods in isolation or simultaneously may lead to significant differences or even reversals in the estimates (see also List, 2002).

Surprisingly, experimental studies measuring consumers' WTP for different products, different specifications, certifications, or informational situations, rarely take into account the influence of the diversity of supply. Many different auction mechanisms (Vickrey second price, random *n*th price, BDM, etc.) and procedures (full biding or endowment) are currently used in these studies. However, whatever the protocol, as soon as many products or rounds are involved, participants are informed from the outset that one binding auction will be randomly selected at the end of the session and that each participant will purchase at most one product. Therefore, some consumers may inflate their WTPs for less-liked products, to avoid ending up buying nothing in case one of those less-liked product would be randomly chosen for the binding sale at the end of the experiment.

This technique may result in biased estimates for two reasons: first, because the consumer is forced to mentally isolate each product offered for sale (since in any case, this isolation will be imposed *ex-post* by the experimenter); second, because this procedure may lead to *ex-post* regret of purchasing a product which does not maximize the consumer's surplus. Moreover, if in many works with experimental auctions the concept of "Consumers surplus" is explicitly used to explain the consumer behaviour or to assess the effects of public policies, it must be noted that, in these experiments, consumers are not confronted with surplus maximization taking into account all the possible purchases (see, for example, Rousu and Corrigan, 2008, Lusk et al., 2005, or Alfnes, 2009).

The endowment approach, in which participants are given one of the tested products and asked to bid for an "upgraded" version, could be viewed as a good way to place consumers in a comparison situation. However, as shown clearly by Alfnes (2009), the endowment approach has the major disadvantage of placing the consumer in an asymmetrical situation where the tested products are not treated equally. This results in well-documented problems (loss aversion, WTP-WTA gap, reference-dependent preferences,…) and also raises the issue that a product is generally not considered as an "upgrade" by all consumers. Moreover, when this method is used to compare more than two products (Roosen *et al.*, 1998; Lusk *et al*., 2004), the random selection of a binding product may again isolate the different alternatives that the experimenter would like to compare.

Choice experiments have the advantage of focusing on comparison between substitutes. However the use of posted prices may strongly influence participants and retrieving individual WTPs for policy simulation is usually not straightforward. That is the reason why we propose to explore another approach taking advantage of experimental auction mechanisms combined with explicit surplus comparison to account for the effect of the diversity of choice alternatives on the decision of buyers. Our method of 'Surplus Comparison Mechanism" (SCM) consists in encouraging consumers to focus on the maximization of surplus and consequently on the comparison of products that are offered on a market. Using a conceptual framework we show, in section 2, how the use of traditional procedures for measuring WTP is not revealing. Conversely, we also show the optimality of SCM that we propose. Then, we propose in sections 3 and 4 an empirical application of the SCM on an example often discussed in the empirical literature on food consumption, *i.e* the WTP for strengthening environmental production conditions of agricultural products (reducing the use of pesticides) and their potential effects on consumer health (reduction of pesticide residues in food). By retaining only three qualitative levels of production (conventional production, Integrated Pest Management, Organic certification), we show how the consumer WTP for one of these three products is strongly dependent, but in a non-erratic manner, on the actual presence on the market of the alternative certifications. It appears that the WTP for one specific certification is decreasing with the number of alternatives presented to consumers and that this result is even stronger when the product is of low quality. The fact that this result is measured structurally on a large majority of the population we have tested, confirms the existence of a "restrictive choice bias" which may be a consequence of decoy and

compromise effects mentioned above. Section 5 is devoted to the conclusion and lessons learned from a political economy perspective.

2. Model and incentive mechanism for revealing WTP

In this section, we propose a model of consumer behaviour, directly derived from Mussa and Rosen (1978) and we justify our approach for measuring WTP in lab experiments. Then we explain the SCM, showing how it is incentive compatible with comparative behaviours for purchasing goods.

2.1 Modelling consumers behaviour with alternative choices

We consider a market represented by J stores indexed by j=1,…,J. Within each store, one or more qualities of the same product are available. We denote by K all possible qualities indexed by k = 1,..., K and it is assumed that the quality of a product is very well marked and recognizable by consumers. Each store j is selling a limited number of possible qualities (subset of K). Denote by $p_{k,j}$ the price of quality k within the store j. There are N consumers, spread into the J stores and indexed by i = 1,..., N. We denote by $U^i(k,j)$ the utility of consumer i (measured monetarily) for purchasing the quality k, in the store j. Note that the introduction of the parameter j in this utility is to capture possible behavioral effects related to a specific salience of quality k for the customer i in the respective store.

Purchase in a single-product store

Consider a store j_0 that has only one type of product quality k_0. A consumer I, present in this store j_0, gets the surplus $S^i(k_0, j_0, p_{k_0,j_0})$ by buying one unit of product:

$$S^i(k_0, j_0, p_{k_0,j_0}) = U^i(k_0, j_0) - \delta^i_{j_0} - p_{k_0,j_0} \quad (1)$$

The parameter $\delta^i_{j_0}$ specifies an "opportunity cost" to purchase in the store j_0. This cost can be positive when it refers to mental conditioning at the time of purchase on the market. In the case of experimental auctions, this can be for example a commitment cost (Lusk et Shogren, 2007, p. 43). The existence of this kind of cost is likely to explain why WTPs are often lower than market prices of the corresponding products (*i.e.* the problem of external validity of lab experiments). The parameter $\delta^i_{j_0}$ could also refer to the unwillingness to purchase the good at the time of experience (this is especially true for perishable or difficult to store food

products). $\delta^i_{j_0}$ could also represent a transportation cost when it is negative (i.e. the opportunity to purchase the good during the time of experience). In this last case we could observe WTPs greater than the prices obtained in the real market. Finally $U^i(k_0, j_0) - \delta^i_{j_0}$ can be viewed as the actual utility of the consumer during the experiment.

Under the conditions of the formula (1), the consumer i buys one unit of the product, if and only if $S^i(k_0, j_0, p_{k_0,j_0}) > 0$, which gives the following condition:

$$p_{k_0,j_0} < w^i_{k_0,j_0} = \text{Max}\{0; U^i(k_0, j_0) - \delta^i_{j_0}\} \quad (2)$$

$w^i_{k_0,j_0}$ represents the willingness-to-pay for the purchase quality k_0 for consumer i in the store j_0. In this case, but only in this case, the consumer is not in a position to compare different possibilities of purchase and has no other option but to buy the product offered to him, if the price is below to its actual utility.

Purchase in a multi-product store

It is now assumed that the consumer is in a store j_1 that offers more than one quality of product (the quality k_0 being supplied with other levels of quality k_1, k_2,…). In this case, the consumer buys the product of quality k_0 if and only if :

$$U^i(k_0, j_1) - \delta^i_{j_1} - p_{k_0,j_1} > \text{Max}\left\{0; \underset{k \neq k_0}{\text{Max}}\{U^i(k, j_1) - \delta^i_{j_1} - p^i_{k,j_1}\}\right\} \quad (3)$$

The willingness to pay for quality k_0 is then the following:

$$w^i_{k_0,j_1} = \text{Max}\left\{0; U^i(k_0, j_1) - \delta^i_{j_1} - \text{Max}\left\{0; \underset{k \neq k_0}{\text{Max}}\{U^i(k, j) - \delta^i_j - p^i_{k,j}\}\right\}\right\} \quad (4)$$

This time the consumer is no longer content to verify that the purchase of the product gives her a positive surplus. She also compares her surplus with the one obtained by purchasing a different product quality, when it is actually available within the store j_1. Thus, the willingness to pay for quality k_1 within the store j_1 can be decreased by the presence of different alternatives of purchase within the store. Consequently the consumer could refuse to buy the quality k_1, even if the selling price of this quality is below its actual utility.

The procedures of choice experiment operate in this way with posted prices which allow to observe the buying behaviour in a situation close to reality. In this case, the number of substitutes, with knowledge of the sales price, can influence the purchase decision and the willingness-to-pay of the product that has to be evaluated.

2.2 Experimental auction and Surplus Comparison Mechanism

An experimental auction is a market where there are a number of products for which we seek to assess the willingness-to-pay of consumers. Therefore the consumer is in a position to compare the products rather than to make independent evaluation regardless of the assessment he can make for the different substitutes. This is the principle of many behavioral economics approaches. This is also the basis of our approach to reveal the WTP in the context of experimental auctions.

In the absence of posted prices it is possible to ask directly to consumers their willingness-to-pay for each of the products and to ensure that such statements is made credible using an incentive mechanism. The most popular mechanisms are Vickrey second price auction, random *n*th price auction and the Becker-DeGroot-Marschak (BDM) mechanism. In all cases, in the absence of behavioral effects related to the salience of products, the willingness-to-pay for one product is not influenced by the presence or absence of substitutes. A sale procedure that uses these mechanisms for each product that is evaluated (or for any of these products randomly drawn) is in principle incentive compatible[3], because the rational choice of the consumer is to reveal his true WTP.

However, there is a strong contradiction between wanting to evaluate the WTP of a product compared to other products (which is routinely done in this type of experiment) and denying consumers the possibility to make comparisons at the moment of the declaration of the WTPs. In fact, if a consumer has a desire to change his willingness-to-pay based on the alternatives proposed in the experimental market, then the use of independent auction mechanisms is not justified. The reason is that if one product is randomly selected, or if each product is sold separately, the consumer is in a buying situation without any possible alternative. He is therefore obliged to reveal a WTP (at the time of his bid) anticipating this sale without possible purchase alternatives. Thus, the stated WTP may not actually reflect the appreciation that he can have for a product compared to alternatives.

[3] Note, however, the criticism of Horowitz (2006), when the distribution of potential sell prices is *ex-ante* known by the consumer.

The Surplus Comparison Mechanism

The surplus comparison mechanism is based on the BDM mechanism with a random price draw for all the products that are tested, and without considering that the distribution of potential market prices is known from the outset by participants. In these conditions it is possible to calculate the consumer surplus, not for a single product selected by the experimenter (or randomly chosen), but for all the products sold during the lab experiment.

Definition (SCM)

Consider N products available on a market. The SCM is defined by a game between a seller and a buyer where the buyer has to propose N bids for buying each product. Then N market prices (one price for each product) are randomly drawn from a distribution unknown to the buyer. The product that gets the highest surplus (difference between the bid and the market price) is sold to the consumer if this surplus is strictly positive. No product is sold to the consumer if this maximum surplus is zero or negative.

Thus, the consumer is well aware that his proposed bid on a given product will be compared to other bids offered on substitutes and that the product she buys will be the one that brings her the highest surplus (and not just a positive surplus).

Note that the SCM is equivalent to the BDM in the case of monoproduct stores (since the condition of positive surplus is equivalent to the condition that the price is below the willingness-to-pay). Note also that it is quite easy to explain this mechanism to consumers in the case of an experimental auction (see the next section below) and that this explanation allows them to focus on the comparison of the different products they have to evaluate.

Proposition

Consider consumers who want to adapt their bids according to the diversity of the products for sale. In a lab experiment, using the BDM mechanism (or any other auction mechanism) on a single alternative randomly chosen among all the products for which consumers bid before the sale, is not incentive compatible. Conversely, the SCM is incentive compatible for revealing willingness-to-pay.

Demonstration

Consider a consumer i and two stores j_1 and j_2 in which are potentially sold two substitutes, k_1 and k_2. Suppose that only the product k_1 is sold in the store j_1 and both products k_1 and k_2 are available in the store j_2. We assume that the consumer WTP for k_1 is different between the two stores, *i.e* $w^i_{k_1,j_1} \neq w^i_{k_1,j_2}$. Let B_1 the bid for quality k_1 of

consumer i in store j_1 and $B_1^{'}$ the bid for quality k_1 of consumer i in store j_2. In accordance with the BDM mechanism it is a dominant strategy for consumer i to bid $B_1 = w^i_{k_1,j_1}$ (because if $B_1 < w^i_{k_1,j_1}$ the consumer could not to buy the product at an acceptable price and if $B_1 > w^i_{k_1,j_1}$ the consumer could have to buy the product at a too high price). In store j_2, if the consumer knows that only one product will be randomly selected, then the optimal bid with the BDM mechanism would be also $B_1^{'} = w^i_{k_1,j_1}$ since the consumer is already confronted with a monoproduct supply, like in store j_1. Consequently the use of BDM mechanism is not incentive compatible. Conversely, with the SCM used in store j_2, suppose the consumer bids a price $B_1^{'} < w^i_{k_1,j_2}$. Then, he takes the risk either to be forced to buy the product k_2, which in reality gives a lower surplus, given the prices drawn at random, or not to buy any products, while the product quality k_2 gave him a strictly positive surplus. If the consumer bids a price $B_1^{'} > w^i_{k_1,j_2}$, he takes the risk to buy the product k_1, while it does not give him the maximum surplus (*i.e* either because he will not be able to buy the product k_2 which would give him the maximum surplus or because he will not be in a position to buy anything if randomly drawn prices are too high for both products in relation to actual WTPs).

To prove experimentally the relevance of this surplus comparison mechanism, two designs can be used: (i) a between-group design in order to show how statistically the value of a product can be significantly different depending on whether or not the product is evaluated alone on the market (ii) or a within-subjects design with different treatments, each one corresponding to a store in which a specific set of products is offered for sale. In this case, if one of these stores is randomly chosen at the end of the experiment (and if the consumer knows that *ex-ante*) it will be possible to measure changes in WTP based on the actual alternative products offered in each store.

Obtaining significant results using a between-group design requires a large number of consumers and a homogeneous population between the different groups. This led us to choose the second option. Admittedly, the within-subjects design, requires more cognitive effort from the consumers, in particular because they had to fully anticipate that *ex-post* they will actually make their purchase in a specific store where some of the products may or may not be offered. However, this can also be viewed as an unfavourable situation to test our predictions.

3. Evidence from the choice between food certifications: protocol

The experiment we have performed to illustrate the relevance of the SCM is linked to the debate on the environmental certification of food products. The number of public and private standards in this sector is particularly important (with different organic certification types and a large number of standards based on Integrated Pest Management, IPM). Therefore consumers face a wide variety of products in the stores' product lines, only on this criterion of differentiation. The market can then be described as a vertically differentiated market with (i) conventional or regular products (without certification), (ii) products with IPM certifications (ii) organic products (Bazoche *et al.*, 2013). It is this classification that we use here to compare the consumers' WTP , *ceteris paribus*, using the SCM.

3.1 Subjects

The experiment was conducted in the sensory laboratory of the *Instituto Superior de Agronomia* (ISA), in Lisbon, with eight sessions. The first session was held in November 2012 (12 participants to test the protocol) and the other seven sessions took place in October 2013 (100 participants who participated in one of 7 sessions). The one hundred and twelve participants were recruited from the general population of Lisbon. A marketing research company was hired to randomly recruit individuals in the age group eighteen to seventy years old to participate in a "fruit and vegetables preference study". Age restrictions were implemented to ensure that a disproportionate number of students or retirees, with relatively low opportunity cost of time, would not dominate the sample. The participants were contacted by phone and they were selected if they ate apples at least once a week and if they regularly participated in the food purchasing. The socio-economic characteristics of the participants were obtained with the recruitment questionnaire. The participants were paid 25€ to participate in the experiment and received their participation fee on arrival at the session when they filled a consent form. One day before the experiment, a convocation letter was sent to all of them with a general explanation of the experiment. In this letter, the random selling-price procedure and the surplus concept were also presented. The number of participants in each session varied from twelve to sixteen persons.

3.2 Products

The products used for this experiment were 1kg bag of 'Golden Delicious' apples obtained from two commercial sources: one supermarket and one organic market, both in Lisbon. The apples were packed into clear bags for easy inspection. Three types of 'Golden Delicious'

apples were defined according to their production method. The first type, "Regular" apple, was presented to the consumers as the apple produced according to the European and National regulation for pesticide use; with no brand, no logo or sticker. The second one, "Integrated Pest Management" (IPM) apple, was presented as the apple type produced with a controlled reduction of pesticide and identified by a national logo (the logo '*Proteção Integrada*'). Finally, the "Organic" apple was presented as the apple type produced without chemical inputs and identified with the European organic label (which is also the logo the most widespread in Portugal for organic products). With the participation of distributors, the three apples certifications were selected with a single size (65-70), with no significant visual differences, including no cosmetic defects.

3.3 Procedure

At the outset of each session, care was taken to make sure all participants were familiar with the concepts mentioned during the experiment. With the support of a presentation, the differences between willingness-to-pay, market price and surplus were explained and different examples were given to facilitate their understanding. At this stage of the presentation, the incentive mechanism, SCM, was introduced. Several examples, with one, two or three products, respectively, were presented to show how the SCM worked. Different purchase and no purchase situations were simulated to emphasize the preference revelation property of this mechanism and to explain why it was in a participant's best interest to bid his true valuation for each product.

Following this first part of the presentation, participants were informed that they would carry out the evaluation of different types of apples in different possible "stores". Participants were also told that a real sale would take place at the end of the session. Specifically, consumers were informed that an urn containing balls on which were inscribed prices will be used to draw market prices for each type of apple (*i.e* for each type of certification) and in every sales situation.[4] At any time during the experiment, consumers had no information about the distribution of prices within the urn. They were only informed that for each participant, one store will be randomly selected at the end of the experiment, and that the market prices will be drawn for each type of certification in this store. Then they will have the obligation to buy one kilogram of the apples that would bring them the highest surplus.

After these explanations and all subsequent questions answered, the experiment started. The three different types of apples were presented to participants at the same time. They were

[4] The urn contained 30 balls, with prices ranging from 0,20€ to 2,00€.

informed about the apples variety ('Golden Delicious') and origin (Portugal), their production methods and the meaning of the two different logos. Also they were previously informed that the three types of apples would not necessary be offered in each store. However, participants were alerted to the fact that at the end of the experiment they would purchase only in one single store.

Seven stores were defined according to all the various possible combinations of different types of apples. During the eight sessions, the order of the presentation of the stores was the same[5], with the following,

Store 1: Selling of the Regular Apple
Store 2: Selling of the Organic Apple
Store 3: Selling of the Regular and Organic Apples
Store 4: Selling of the IPM Apple
Store 5: Selling of the Regular and IPM Apples
Store 6: Selling of the IPM and Organic Apples
Store 7: Selling of the Regular, IPM and Organic Apples

Participants did not know in advance the series of stores they would go through. During each step, apples were presented and evaluated simultaneously. For all the steps, the participants could only make a visual inspection of the products and examine the labels. They could not taste the apples. A maximum purchase price for each apple type was recorded for each participant at the end of each step. Participants were informed that if they did not like a specific type of apple, they always had the possibility to indicate a buying price of zero. When a new step started, participants could not go back or change the prices they had given.

During the last phase of the experiment, the prices of different types of apple, available in each store, were randomly drawn from the urn. These prices were written on a whiteboard to make them visible to the participants. After that, each participant randomly drew a retail store (Selecting a numbered card from 1 to 7). In this store and following the SCM procedure, the participant' bids were compared with the market prices, and the participant had to buy 1 kilo of the apple type that gave him the highest surplus.

[5] Usually the order of presentation plays an important role given surprise effects that consumers can have during an experiment. However in our case, consumers were fully informed of potentially available products in a store. This limited the effect of the presentation order of the choice set.

4. Results

4.1 Sample and Data

As the recruitment procedure and protocol were the same, we decided to pool the data of the eight sessions. After some control for coherence of the data, we decided to exclude one participant of the October 2013 sessions, whose mean WTP was 4.75 times higher than the mean WTP of all other participants and 81% higher than the mean WTP of the second highest bidder of the sample. So we ended up with a total of 111 participants for the analysis. Table 1 shows the main socio-demographic characteristics of the sample.

Table 1. Sociodemographic characteristics of the sample

	Mean	S.D.	Min	Max
Number of participants	111			
Number of women	60			
Number of men	51			
Age	40	14	18	65
Education (1=primary; 2=secondary; 3=some college or more)	2.5	0.55	1	3
Income (€/month per capita)	840	561	150	3250
Household size	2.6	1.29	1	8
Number of children	0.45	0.82	0	4

The 111 participants included in our analysis valued 3 types of apples (regular, IPM, and Organic) in 7 different stores, corresponding to as many different choice sets[6] (3 choice sets with only one apple, 3 with 2 apples, and 1 with 3 apples. Each apple was then evaluated 4 times (once alone, twice with one of the other two variants and once with the other two variants), which means 12 evaluations for each participant and a total of 1332 WTPs. The mean and median WTP are equal to 1.2€/kg.

Almost 10% (132) of WTPs are equal to zero, but no participant refused to buy systematically. The maximum number of refusals per participant is 8, the average number is 1.1 and the median is 0.

4.2 WTP for regular, IPM and organic apples according to the choice set

The main empirical issue of the paper is to test whether WTP for a given product is affected by the choice set taken into consideration when consumers are assessing the values of the different products. It is important to stress again that the point of interest, here, is the size and

[6] Throughout the rest of the paper the term "store", which was supposedly more familiar to consumers, will be replaced by "choice set", which is conceptually more relevant. This will allow us to rearrange the order of the choice sets, to focus first on those that refer to single-product situations (benchmarks of our study).

composition of the choice set and not the information related to each of the alternatives presented. When participants started their evaluations of the products in the different choice sets they already had all the information about the characteristics of the different products, and no further information was released until the end of the experiment.

Table 2 shows mean prices (and corresponding standard deviations of the mean) for each of the three apples according to the choice set considered when the assessment was made. When considered in isolation, in choice sets 1 to 3, WTP are 1.16€/kg for the conventional apple, 1.23€ for the IPM apple, and 1.54 for the organic. As can be seen directly from Table 2, WTPs are lower for all apples as soon as they are considered jointly with one or two other variants in choice sets 4 to 7.

Table 2. Mean WTP according to apple and choice set

	Choice set							
Apple	1	2	3	4	5	6	7	Mean WTP
Organic (€/kg) (SE of mean)			1.54 (0.06)		1.48 (0.06)	1.43 (0.06)	1.44 (0.06)	1.47 (0.03)
IPM (€/kg) (SE of mean)		1.23 (0.05)		1.18 (0.06)	1.17 (0.06)		1.15 (0.05)	1.18 (0.03)
Regular (€/kg) (SE of mean)	1.16 (0.05)			0.96 (0.06)		0.97 (0.06)	0.88 (0.06)	0.99 (0.03)
Mean WTP (€/kg) (SE of mean)	1.16 (0.05)	1.23 (0.05)	1.54 (0.06)	1.07 (0.04)	1.32 (0.04)	1.20 (0.05)	1.16 (0.03)	1.22 (0.02)

To test whether these changes in WTP are significant, two linear regressions of individual WTP on apple type and choice set have been run. Results are displayed in Table 3. The first regression (1) measures the impact of apple type when apples are evaluated alone (choice sets 1 to 3, variables Organic, IPM and Regular) and when they are evaluated jointly with one other apple or all together (choice sets 4 to 7, variables OrganicXAlter, IPMXAlter and RegularXAlter). The IPM apple evaluated alone (in choice set 2) is the reference. Results show that when evaluations are made separately, the organic apple is valued higher than the other two apples (IPM and Regular), for which WTPs are not significantly different. When evaluated jointly, WTPs for all three apple decrease significantly compared to WTPs evaluated separately. Testing the differences between the three pairs of coefficients shows that

the decrease in WTP is more pronounced for the regular apple compared to the IPM and Organic, for which the decrease is not significantly different.

Table 3. Impact of choice set on WTP

	(1) Apple and choice set		(2) Apple and choice set size	
	Coefficients	Standard error	Coefficients	Standard error
Organic	0.308***	0.051	0.292***	0.045
IPM	ref	ref	ref	ref
Regular	-0.074	0.039	-0.190***	0.039
OrganicXAlter	-0.086**	0.025		
IPMXAlter	-0.064**	0.019		
RegularXAlter	-0.220***	0.003		
Choice set size = 1			ref	ref
Choice set size = 2			-0.110***	0.016
Choice set size = 3			-0.151***	0.020
Constant	1.230***	0.055	1.275***	0.053
Observations	1332		1332	

Legend: * p<0.05; ** p<0.01; *** p<0.001
Standard errors are adjusted to account for intra-participant correlation.

The second model regresses WTP on apple type and choice set size separately, again the IPM is the reference for apple type and size =1 for choice set size. As shown in Table 3, third column, all coefficients are highly significant. Testing for equality shows that they also significantly differ between apples (Organic vs. Regular, $p<0.0001$), and between choice set size (size = 2 vs. size = 3, $p=0.006$). This shows clearly that, when choice set size is accounted for, WTPs for the three apples are significantly different. Moreover, it is also clear that mean WTP decreases (by 0.11€ and then again by 0.04€) when the size of the choice set increases from one to three.

To sum up, these results, show that:

- When apples are evaluated in isolation, WTP for IPM and Regular apples are not significantly different, but when the choice set size is accounted for, WTP for all three apples appear to be significantly different.
- When apples are evaluated jointly (2 by 2, or all together), WTP is lower, and WTP decreases when choice set size increases. The decrease being more pronounced for the least valued apple (the conventional).

4.3 Individual behaviours

Considering different alternative products jointly instead of separately leads to a decrease in the mean WTP for each alternative. This could result from different individual behaviours. For example, in a situation where they have access to a close, and possibly more attractive, substitute, consumers may refuse to buy one of the products ("boycott" effect) or reduce the price they offer to pay for this product. Some consumers may also be insensitive to the choice set, and stick to their evaluation of each product whatever the substitutes available at the moment of purchase. Moreover, we cannot exclude that the availability of a substitute may increase the WTP for a product evaluated alone, if the presence of the substitute makes the product look better than when it is considered alone. Which behaviour will prevail is an empirical question, depending on individuals, products and choice sets. To investigate these different possibilities, we have compared WTP for each participant and each apple, when the apple is evaluated alone and when it is evaluated jointly with one or two alternatives. For the sake of simplicity, we comment only on this last case, that is when the three products are evaluated together compared to separately (the detailed results for all possible cases are displayed in the annex at the end of the paper).

When the choice set changes, participants can stick to their previous evaluation, decrease or increase their WTP. Previous results indicate that increasing WTP is a marginal reaction and that decreasing WTP is much more frequent. A mean decrease in WTP may be the outcome of two different responses: a consumer may stop buying (then WTP drops to zero) or keep on buying, but at a lower price. Both reactions are worth measuring as they may have different impacts on the market.

In Table 4, participants are distributed according to their change in WTP for each apple, when evaluations are made separately compared to when they are made jointly for the three apples. Four cases are presented: no change, increase in WTP, stop buying, decrease in WTP. Below the number of subjects corresponding to each case, average WTP under both conditions (alone/together) are shown, as well as the impact on the mean WTP for each apple.

Table 4. Change in average WTP when evaluations are made separately or jointly

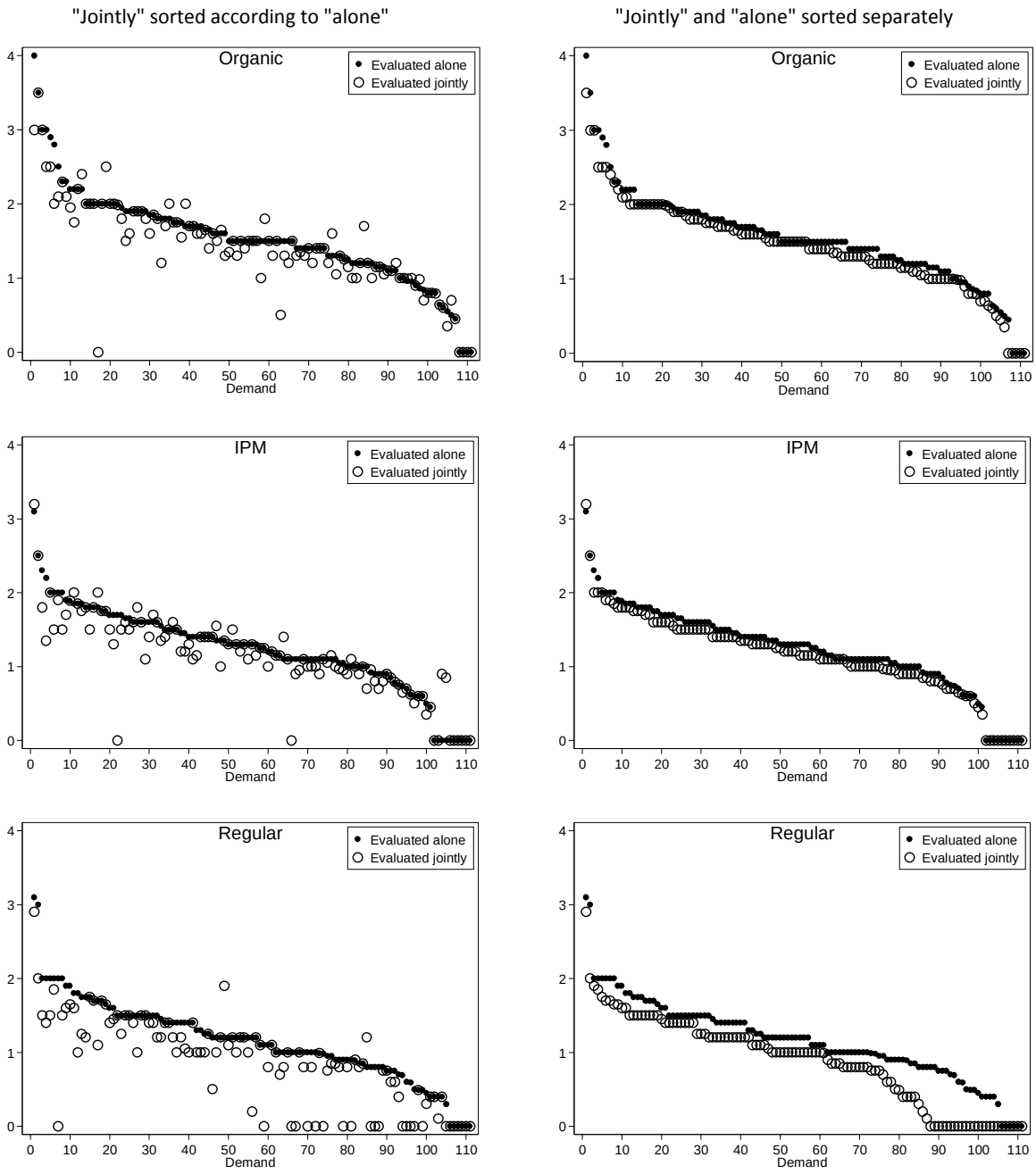
	Change in WTP according to choice set size				
	Stable	**Increase**	**Stop buying**	**Decrease**	**Total**
Organic					
Number of subjects	54	13	1	43	111
WTP if evaluated alone (€/kg)	1.44	1.36	2.00	1.71	1.54
WTP if evaluated together (€/kg)	1.44	1.58	0.00	1.44	1.44
Impact on average WTP (€/kg)	0.00	0.02	-0.02	-0.10	-0.10
IPM					
Number of subjects	48	15	2	46	111
WTP if evaluated alone (€/kg)	1.08	1.27	1.40	1.36	1.23
WTP if evaluated together (€/kg)	1.08	1.50	0.00	1.15	1.15
Impact on average WTP (€/kg)	0.00	0.03	-0.02	-0.09	-0.08
Regular					
Number of subjects	36	3	18	54	111
WTP if evaluated alone (€/kg)	0.98	0.96	0.86	1.39	1.16
WTP if evaluated together (€/kg)	0.98	1.33	0.00	1.08	0.88
Impact on average WTP (€/kg)	0.00	0.01	-0.14	-0.15	-0.28

Results are very similar for the Organic and IPM apples, both differ from the regular. The conventional apple is clearly the most affected when it is evaluated jointly with the other two apples. Almost 50% of the participants (54) decrease their willingness to pay and 18 are not willing to buy anymore. Although different in size (18 vs. 54), these two groups have almost the same impact (-0.14 and -0.15€/kg) on the average decrease of the WTP for the regular apple (-0.28€/kg). Roughly, one out of three participants (36) simply did not change their WTP, and increasing it appears anecdotal. Things are different for the organic and IPM apples, for which the most frequent response to a change in the choice set is to do nothing. The second most frequent response is a decrease in WTP. Almost no participant has chosen to stop buying. Interestingly, increases in WTP may happen (13 cases for the organic apple and 15 for the IPM), which suggests that some participants have considered theses apples more favourably when they were confronted to substitutes. Finally, it is worth noting that, for these two apples, the final impact of the choice set on the average WTP is almost entirely due to a decrease in WTP.

The framing of the evaluation of individual WTP may impact the simulation of market shares or demand functions used to build marketing or public policy scenarios. To have a rough idea of this impact, cumulative distributions of WTP for each apple are shown in Figure 1. Black dots are for WTP evaluated separately for each apple, and hollow circles are used when evaluations are made jointly for the three apples. In the left column, individuals are ranked according to their WTP when apples are evaluated alone. Individual responses appear

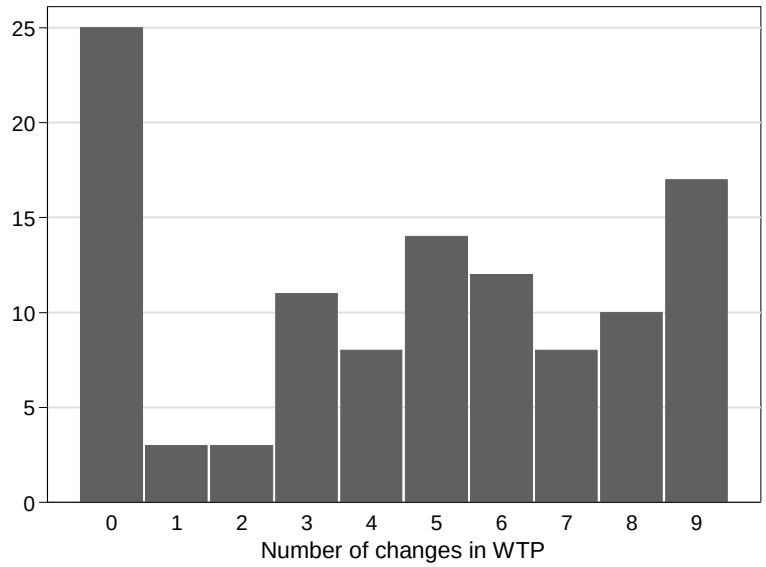
quite heterogeneous. The main impact of assessing WTP jointly is a decrease in WTP, clearly more frequent for the regular apple. In the right column, cumulative distributions are drawn independently. This shows that demand curves drawn from joint evaluations are below demand curves drawn from separate elicitation of WTP. Again, the difference between the two estimations is much more pronounced for the regular apple than for the IPM and Organic, for which demand curves are very similar under both conditions.

Figure 1. Cumulative distribution of WTP when apples are evaluated alone or jointly



Previous results have shown a diversity of responses to changes in choice set's size. A last question is to know whether this diversity is between or within individuals. Measuring individual sensitivity to differences in choice framing may bring some insight into that question. Each individual evaluated each apple 4 times: once alone, twice with one other apple and once with the two other apples. Comparing WTP of one apple evaluated alone to WTP when the apple is evaluated jointly with other apples, produces 3 pairs of WTPs per apple, thus a total of 9 pairs for each participant. The two WTPs of each pair can be identical or different, which allows to compute the total number of changes in WTP for each participant, ranging from 0 (same WTP for an apple whatever the choice set) to 9 (WTP changes each time the choice set is different). Figure 2 shows the distribution of the 111 participants according to the number of changes in WTP.

Figure 2. Distribution of participants according to their number of changes in WTP



It appears that systematic unresponsiveness to choice set is the behaviour of a significant minority: 25 participants (22.5%) never changed their WTPs. These subjects have a given WTP for each apple (none of them had the same WTP for the 3 apples), and they never change it. This does not mean that they do not discriminate between apples: their WTPs are different and increasing from the regular to the IPM and the organic. Conversely, a significant number of participants (17) change their WTP at each change of the choice set. The remaining 69 participants show some sensitivity to the size of the choice set, most of them change their WTP systematically for one apple (3 changes) or for two (6 changes). Socio-demographic

characteristics of participants have no significant influence on the number of changes, except age (p=0.03), older participants being less sensitive to the choice framing.

5. Concluding Comments

In this paper we proposed a new WTP elicitation procedure based on the BDM mechanism and allowing for consumer surplus maximisation. This Surplus Comparison Mechanism (SCM) has the advantage of focusing on comparative effects in assessing the purchasing behaviour of consumers when several products are tested simultaneously. Our empirical work provides original results that justify the use of this mechanism, while showing the operational validity for lab experiments. This experimental design, which main objective was to validate the methodology, requested a significant cognitive effort, going beyond what consumers usually do when they choose a product in a real market. Indeed, the exploration of all possible choice situations during the experiment, and the need for consumers to mentally isolate these choice sets, were admittedly demanding. However, the main result of this lab experiment (*i.e.* the decrease of WTP according to the number of substitutes within the choice set) appears to be significant, in particular when considering the fact that consumers were *ex-ante* fully informed on the products potentially available on the market. Therefore, there was no element of surprise when each choice set was proposed to the consumers (the surprise effect that often explains the decrease of WTP for conventional products when innovative products reach the market or appear in an experiment).

This validation experiment for SCM should now be replicated using a between-group design. Therefore participants would not have to focus on the random potential draw of each market situation at the end of the session of the lab experiment. There is little doubt that the decrease in the average WTP, based on the number of substitutes, would be much higher than the one we obtained in this paper. This experiment could be done using the case studies we presented in the preamble of this article, for example taking the idea that GMO products would obtain a much higher WTP if GMO free products were not available on the market.

These results have implications for the development of public policy. If we return to the example of the labeling or the ban on the use of GMOs (Noussair et al, 2004; Lusk et al, 2005 for previous studies) it is clear that the boycott observed on many markets for these products depends a lot on the presence of GMO free products. Hence public policy in some countries could be reconsidered: in terms of barriers to trade, market bans or food products labeling, to both denounce the presence of GMOs or certify GMO free in Europe and the United States.

Considering the scope of this paper, the key challenge for certified products on environmental practices lies in strengthening or not the Minimum Quality Standards. Since the conventional production would be banned from the market to force the adoption of IPM by farmers, the consequences in terms of consumer surplus could be largely modified, and in this case accounting for the actual choice set in the evaluation of WTP is clearly a major issue. If, as suggested by our study, the eviction of conventional production rather caused an increase in WTP for IPM and organic products, there would be a trade-off to make with the direct effects of product diversity.

Appendix.

Number of Participants and WTP by apple and choice set according to change in WTP when the choice set size increases

	Apple and choice set								
	Organic			IPM			Regular		
Subjects and WTP	5	6	7	4	5	7	4	6	7
No change									
Number of subjects (1)	4	4	4	8	9	8	6	6	6
WTP if no alternative (€/kg)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
if 1 or 2 alternatives (€/kg)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Number of subjects (2)	57	58	50	60	59	40	39	46	30
WTP if no alternative (€/kg)	1.52	1.56	1.55	1.31	1.36	1.30	1.19	1.22	1.17
if 1 or 2 alternatives (€/kg)	1.52	1.56	1.55	1.31	1.36	1.30	1.19	1.22	1.17
Decrease in WTP									
Number of subjects (3)	1	3	1	3	3	2	14	12	18
WTP if no alternative (€/kg)	2.00	1.45	2.00	1.48	1.32	1.40	0.82	0.93	0.86
if 1 or 2 alternatives (€/kg)	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Number of subjects (4)	32	37	43	27	28	46	49	44	54
WTP if no alternative (€/kg)	1.76	1.70	1.71	1.40	1.28	1.36	1.37	1.32	1.39
if 1 or 2 alternatives (€/kg)	1.51	1.46	1.44	1.26	1.08	1.15	1.15	1.09	1.08
Increase in WTP									
Number of subjects (5)				2	1	2			
WTP if no alternative (€/kg)				0.00	0.00	0.00			
if 1 or 2 alternatives (€/kg)				0.85	0.99	0.88			
Number of subjects (6)	17	9	13	11	11	13	3	3	3
WTP if no alternative (€/kg)	1.53	1.43	1.36	1.44	1.53	1.46	1.07	1.00	0.96
if 1 or 2 alternatives (€/kg)	1.74	1.63	1.58	1.53	1.65	1.60	1.28	1.20	1.33
Total									
Number of subjects	111	111	111	111	111	111	111	111	111
WTP if no alternative (€/kg)	1.54	1.54	1.54	1.23	1.23	1.23	1.16	1.16	1.16
if 1 or 2 alternatives (€/kg)	1.48	1.43	1.44	1.18	1.17	1.15	0.96	0.97	0.88

(1) Number of subjects with WTP=0 for a given apple when there is no alternative, and WTP=0 when there are alternatives.
(2) Number of subjects with WTP>0 for a given apple when there is no alternative, and constant when there are alternatives.
(3) Number of subjects with WTP>0 for a given apple when there is no alternative, and WTP=0 when there are alternatives.
(4) Number of subjects with WTP>0 for a given apple when there is no alternative, and lower when there are alternatives.
(5) Number of subjects with WTP=0 for a given apple when there is no alternative, and WTP>0 when there are alternatives.
(6) Number of subjects with WTP>0 for a given apple when there is no alternative, and higher when there are alternatives.

References

Alfnes, F. (2009), "Valuing product attributes in Vickrey auctions when market substitutes are available", *European Review of Agricultural Economics*, Vol 36 (2), 133-149.

Ariely, D. (2008) "Predictably Irrational: The Hidden Forces That Shape Our Decisions*"*, New York, NY: Harper Collins Publishers.

Bazerman, M., Moore, D., Tenbrunsel, A., Wade-Benzoni, K., Blount, S. (1999), "Explaining how preferences change across joint versus separate evaluation", *Journal of Economic Behavior & Organization*, Vol. 39 (1999) 41–58.

Bazoche, P., Combris, P., Giraud-Héraud, E., Seabra Pinto , A. , Bunte, F. Tsakiridou, E. (2013), "Willingness to pay for pesticide reduction in the EU: nothing but organic?", *European Review of Agricultural Economics*, 1–23.

Bordalo, P., Gennaioli, N., Shleifer, A. (2013), "Salience and Consumer Choice," *Journal of Political Economy*, University of Chicago Press, vol. 121(5), 803 - 843.

Cohen J., Basu, K. (1987), "Alternative Models of Categorization : Toward a contingent Processing Framework", *Journal of Consumer Research*, vol 13.

Horowitz, J.K. (2006), "The Becker-DeGroot-Marschak mechanism is not necessarily incentive compatible, even for non-random goods", *Economic Letters*, 93, 6-11.

Huber, J., Payne, J.W., Puto, C. (1982), "Adding asymmetrically dominated alternatives: Violations of regularity and the similarity hypothesis", *The Journal of consumer research*, 9 (1), 90

Kahneman, D., Tversky, A. (1979), "Prospect theory: An analysis of decision under risk", Econometrica, 47, 263-291.

List, J. (2002), "Preferences Reversals of a Different Kind: The "More Is Less" Phenomenon", *The American Economic Review*, vol 92, N°5, 1636-1643.

Lusk, J., House, L., Valli, C., Jaeger, S., Moore, M., Morrow, B., Traill, W.B. (2005) "Consumer welfare effects of introducing and labelling genetically modified food", *Economics Letters*, 88, 382–388.

Lusk, J.L., Feldkamp, T., and Schroeder, T., (2004), "Experimental Auction Procedure: Impact on Valuation of Quality Differentiated Goods", *American Journal of Agricultural Economics*, 86 (2), 389-405.

Lusk, J.L., Jamal, M., Kurlander, L., Roucan, M., Taulman, L. (2005) "A meta-analysis of genetically modified food valuation studies", *Journal of Agricultural and Resource Economics*, 30, 28–44.

Mussa, M, Rosen, S. (1978), "Monopoly and product quality", *Journal of Economic Theory*, Volume 18, Issue 2, August 1978, 301–317.

Noussair, C., Robin, S., Ruffieux, B. (2004), "Do consumers really refuse to buy genetically modified food?", *The Economic Journal*,114, 102–120.

Pothos, E.M., Wills, A.J. (2011), "Formal approaches in categorization", Cambridge University Press, 308p.

Roosen, J., Fox, J.A., Hennessy, D.A., Schreiber, A., (1998), "Consumers' valuation of insecticide use restrictions: an application to apples", *Journal of Agricultural and Resource Economics*, 23(2): 367-384

Rousu, M., Corrigan, J. (2008), "Estimating the Welfare Loss to Consumers When Food Labels Do Not Adequately Inform: An Application to Fair Trade Certification, *Journal of Agricultural & Food Industrial Organization, Volume* 6, *Article* 3.

Simonson, I., (1989), "Choice based on reasons: The case of attraction and compromise effects ", *The Journal of consumer research*, 16 (2), 158

Tversky, A. and Simonson, I., (1993), "Context-Dependent Preferences", *Management Science*, 39 (10), 1179.