JAE Journal of Agricultural Economics

Journal of Agricultural Economics, Vol. 71, No. 3, 2020, 778–797
doi: 10.1111/1477-9552.12376

Labels for a Local Food Speciality Product: The Case of Saffron

Ana I. Sanjuán-López and Helena Resano-Ezcaray

(Original submitted June 2019, revision received December 2019, accepted January 2020.)

Abstract

We examine the potential demand for a local food speciality product, saffron, with alternative labels, using a choice experiment. The paper contributes to the literature on credence attributes, by examining Willingness to Pay (WTP) for the local, organic and PDO (Protected Designation of Origin), their differences across experimental conditions (hypothetical and non-hypothetical), and by identifying the effects of personal characteristics, in terms of socio-demographics and level of product involvement on the differences in WTP. We find that the local saffron speciality has an important appeal that could be better reinforced with the PDO rather than the organic labelling, and that consumers show a consistent pattern of preferences across experimental environments. WTP tends to be higher in the hypothetical setting and, in particular, consumers with relatively more knowledge and deeper roots in the territory tend to exhibit a larger WTP premium for local origin and its certification. These results may help producers improve their marketing of agri-food products with a high gastronomic value and differentiation potential, while they warn about an overstatement of WTP for socially desirable characteristics, such as organic labelling, which is also relevant for policymakers.

Keywords: *Credence attributes*; *hypothetical and real choice experiments*; *labelling*; *sustainability*; *Willingness to Pay*.

JEL classifications: *D12, Q13, Q18, Q56*.

Ana Sanjuán-López is a Researcher at the Agrofood Research Centre of Aragón (Instituto Agroalimentario de Aragón, IA2-CITA-Universidad de Zaragoza), Zaragoza, Spain. E-mail: aisanjuan@aragon.es for correspondence. Helena Resano is a lecturer at the University of Zaragoza, Department of Agricultural Sciences and Natural Resources, Zaragoza, Spain (Instituto Agroalimentario de Aragón, IA2-CITA-Universidad de Zaragoza). The authors gratefully acknowledge financial support from the National Institute of Agrofood Research and Technology (INIA) of Spain through the project PET2007-14-C05-04. The authors are grateful to the anonymous referees and the Editor David Harvey, for their constructive comments and thorough revision.

1. Introduction

Throughout history, Spain has been an important saffron producer, although recently Spanish saffron production has struggled to compete with cheaper sources from the Middle East (principally Iran). Saffron is a highly labour-intensive crop and higher unit labour costs coupled with rural depopulation have led to a progressive abandonment of saffron cultivation. Spanish production has declined from 124 t at the beginning of the 20th century (Ávila, 1999) to 1,345 kg in 2007 (MARM, 2010). Private and public initiatives to revitalise the crop, under the auspices of the EU rural development policy, coupled with the economic crisis, have contributed to a relative boost of saffron cultivation, and the most recent data indicate production of 1,973 kg in 2016 (MARM, 2017). Currently, Spanish saffron production is largely restricted to the region of Castilla La Mancha, and to a lesser extent, the province of Teruel (around 3.5%) in the North-Eastern region of Aragón. Despite such unfavourable economic factors, Spanish saffron still enjoys a reputation of high quality, whilst the European Protected Designation of Origin (PDO) for Castilla La Mancha, in place since 2001, has potentially reinforced its quality status.

The province of Teruel is predominantly rural, sparsely populated, and is classified as a 'less favoured area' in the EU rural development schemes (PDR, 2016). In this context, it is important to find alternative or complementary sources of income that contribute to economic and consequently social sustainability, and saffron could be one of those sources. Apart from quality reputation, Teruel saffron may also benefit from a combination of factors with increasing resonance amongst consumers, such as being local, sustainable or organic. The local attribute implies reduced distance and associated perceived environmental benefits (Marsden and Smith, 2005) as well as social responsibility, as the purchase of local food promotes local businesses and jobs (Toler *et al*., 2009; Denver and Jensen, 2014), which in turn, may be reinforced when driven by ethnocentrism (Sanjuán-López *et al*., 2009). The production of saffron in Spain is still a low-input production system (Sanjuán-López *et al*., 2009), where conversion to organic status is relatively straightforward, as traditionally no synthesised fertilizers are used. From a technical point of view, experiments applied to Teruel saffron have revealed specific traits attributable to soil and climate, as well as the drying process (Alvarez and Mallor, 2009), establishing a basis for origin certification. Protected Designation of Origin (PDO) and Protected Geographical Indication (PGI) are the most relevant EU labelling schemes in terms of product and countries coverage. Using a price hedonic function, Sanjuán-López *et al*. (2009) estimated that the only existing Spanish PDO currently in the market (PDO La Mancha) adds an extra 10% value to the already highly valued La Mancha origin.

Saffron is mainly used as a condiment that provides colour, smell and taste. Nevertheless, saffron can also become an ingredient in innovative food products, adding a distinctive value (Sanjuán-López *et al*., 2011). Historically, saffron has also been used for health (Kyriakoudi *et al*., 2015) and cosmetic purposes (Bathaie and Mousavi, 2010). Nevertheless, the pharmacological industry is still in its infancy, and the use in cosmetics at an industrial scale is only partially developed.

Improving the commercialisation in local markets might help to trigger production and therefore contribute to arresting the long-term decline in this sector, generating income opportunities that may help the preservation of rural likelihoods. Currently, Teruel saffron is marketed through two private brands with only a small share being

designated organic. We use a choice experiment to explore the market value of these potentially important characteristics.

We also conduct both a stated or hypothetical (SCE) and a real or non-hypothetical choice experiment (RCE) to explore and mitigate possible hypothetical bias. A few studies in the food literature have compared both experiments, and although mixed results are found, in general, an overstatement of WTP in hypothetical settings predominates. Thus, Lusk and Schroeder (2004) in their study on beef, find a significantly higher total WTP in the hypothetical than in the non-hypothetical choice experiment (i.e. product defined by a combination of attributes and levels), although not in marginal WTP for the quality label. Similarly, Alfnes and Steine (2005) find a significant upward hypothetical bias in total WTP for salmon. Finally, both Yue and Tong (2009) and De Magistris *et al*. (2013) find an upward hypothetical bias in marginal WTP for the organic and local attributes in fresh tomatoes and almonds, respectively. Assuming the presence of such an hypothetical bias, additional relevant literature has opted directly for non-hypothetical or incentive-compatible experiments, such as Alfnes *et al*. (2006), for the colour of salmon; Olesen *et al*. (2010), for organic and welfare certified salmon; Gracia (2014) for local lamb; and De Magistris and Gracia (2016) for almonds.

Our aim is twofold. First, we investigate consumers' preferences and WTP for those attributes that could help to improve the marketing of Spanish saffron (local origin, the PDO label and the organic production process). Second, we investigate whether WTP differs between hypothetical and non-hypothetical conditions, and try to identify particular consumer profiles (in terms of socio-demographics and product involvement) which are more prone to hypothetical bias. Our overall objective is to provide accurate and reliable recommendations to actors involved in the saffron marketing chain.

Our analysis contributes empirically to three strands of literature. First, there is an ongoing debate in the sustainable food consumption literature about whether the 'local' versus the 'organic' claim provides a higher utility to consumers (e.g. Denver and Jensen, 2014; Gracia *et al*., 2014; De magistris and gracia, 2016). Second, we provide further evidence about the existence (or not) of an upward hypothetical bias in choice experiments. While there is a general urge to understand the role that consumers' characteristics play in food choice (Bernués *et al*., 2003; Resano *et al*., 2018; Peschel *et al*., 2019), to the authors' knowledge, little is known about how these characteristics influence hypothetical bias. Relevant exceptions consider psychographic characteristics such as agency and openness (Grebitus *et al*., 2013), or consumer-related characteristics and motives (e.g. product category knowledge) (Hofstetter *et al*., 2013) as drivers of the hypothetical bias. Third, we illustrate the use of individual Status Quo (SQ), a novel feature in the design of stated choice experiments that allows a closer representation of the choice alternatives. Individually specific SQ is highly recommended to avoid the loss of information or misinterpretation of the most common use of 'non-chosen' or 'opt-out' alternatives to prevent forced choices (e.g. indifference between levelled alternatives, disinterest, escape from cognitive burden, etc.) (Hensher *et al*., 2015), which in turn poses challenges in the econometric estimation and potential biased estimates (Pedersen and Gyrd-Hansen, 2013). To the authors' knowledge, there are no previous examples of choice experiments on food choices with individual Status Quo alternatives.

The rest of the paper is organised as follows. The next section describes the experiment stages, the design of the hypothetical and non-hypothetical choice experiments,

the econometric modelling and statistical tests. Section 3 describes the results and Section 4 concludes.

2. Materials and Methods

2.1. Experiment stages

The experiment took place in the main city of the region of Aragón (Zaragoza), the region where the Teruel saffron is produced, between June–July 2009.[1] The sample includes 208 consumers, stratified according to gender and age. Consumers were randomly recruited through an advertisement placed at civic centres and citizens associations. The target population comprised consumers older than 18 years old and responsible for food shopping.

Table 1 presents the main socio-demographic characteristics of the sample. Compared with relevant regional/local population statistics, our sample is biased toward women who, in Spain, are still primarily responsible for household food shopping and preparation. There is a moderate over-representation of younger and more educated respondents. This is commonplace in market research, as these subjects usually demonstrate a greater willingness to participate in experimental studies of this nature. On the other hand, a younger sample is to be considered as an advantage when assessing potential markets. There is a moderate over-representation of the middle-income interval (although not statistically significant) and traditional household composition (i.e. two adults plus children).

Face-to-face experiments were conducted in small groups of no more than 12 participants (8 participants on average). First, the stated choice experiment was conducted, where a prior cheap talk script was provided to reinforce the idea that choices needed to reflect what participants would buy following their preferences and being fully aware of their budgetary constraints (e.g. Silva *et al*., 2011). After 15 minutes devoted to some additional tasks including a questionnaire collecting socio-demographic characteristics, purchase and consumption habits, we then implemented a real or non-hypothetical experiment.

In the real experiment, participants were offered 5 euros to buy saffron. They had to move to a different place in the room, where packages of half a gram were displayed with its corresponding identifying labels (as in Figure S1 in the online Appendix). The selection of the binding scenario followed Alfnes *et al*. (2006). That is, at the end of the choices, each respondent chose blindly a choice set and had to buy the chosen alternative in that choice set. The same choice sets were used in the stated and real experiments, and the same respondent faced the same block of choice sets in both experimental conditions. Thus, a within-subject experiment was carried out, which allowed analysis with more powerful econometric techniques, which is especially suitable for testing consumers' utility theory (Charness *et al*., 2012). Nevertheless, and in order to mitigate a possible carry-over effect inherent to a within-subject experiment, both temporal and spatial distance was provided for our respondents' SP and RP choices, as suggested by Charness *et al*. (2012).

[1]A recent follow-up of the study by Sánchez *et al*. (2018) points at no-significant differences in terms of consumption and purchase habits of saffron, see Section 3.1.

Table 1
Description of the sample

Variable	Indicator	% sample	% population
Gender*,†	Male	36.1	48.6
	Female	63.9	51.4
Age†	≤34 years old	30.3	27.9
	35–44 years old	18.8	20.0
	45–54 years old	23.1	17.0
	≥55 years old	27.9	35.1
Household size‡	—	3.1 persons	2.6 persons
Household composition*,‡	Unipersonal	9.7	18.3
	Two adults without children	25.5	29.4
	Two adults with children	42.3	26.1
	Others	22.4	23.5
Education***,‡	Compulsory school (primary, secondary)	26.1	50.6
	High school or equivalent	28.0	23.3
	University	45.9	26.1
Net household income (€/month)‡	≤1,500 €/month	30.5	38.0
	1,500 €–3,000€/month	49.8	37.0
	≥3,000 €/month	19.7	25.0

Notes: *, ***Mean significant differences at 10% and 1%, respectively, between the sample and population relative frequencies according to the chi-square test.
†Percentages over the population of the city of Zaragoza in 2010.
Source: IAEST (2010a).
‡Percentages over the population of Aragón in 2009.
Source: IAEST (2010b, 2011).

2.2. *Stated and real choice experiments*

A choice experiment was used to elicit consumers' preferences towards saffron, as it allows us to deal simultaneously with an array of attributes, mimicking a real purchase situation, while from a theoretical perspective, it is consistent with the Random Utility Model (Louviere *et al*., 2000). Choice experiments have become an important valuation tool in food marketing to evaluate non-existent products or potential innovations in food (Bazzani *et al*., 2017). Choices can be elicited in hypothetical (Stated Choice Experiment – SCE) or non-hypothetical (Real Choice Experiment – RCE) experimental conditions. In contrast with the former, a non-hypothetical or real choice experiment accompanies the description of the product (e.g. presented in cards) with the real product and makes the experiment's economic incentive compatible, i.e. following some random selection procedure (see below) the respondent is bounded by the choices he/she made and has to proceed to the actual purchase of the product selected.

Consideration of available saffron and substitute products in the market informed the specific design of the choice experiment. In our experiment, the consumer had to choose between the traditional 'strand' and the convenient 'ground' options. Specific prices for 'strand' and 'ground' were considered to reflect the range of market prices (€1.5, €2.25 and €3 per half gram for saffron in strands, and €1, €1.75 and €2.5 per

half a gram for ground saffron). The remaining attributes were: (i) Origin: 'La Mancha' (main producing area in Spain); 'Teruel' (the local producing area); 'Iran' (owing to its importance in Spanish imports and international markets); (ii) Quality Certification: 'PDO' or 'None'; and (iii) Production method 'organic' or 'conventional'. Brief explanations of the meaning of organic production and the PDO label were given to the participants before both experiments.

We proposed a 'labelled' experiment as price and PDO are specific to each format of the product, i.e. 'strand' average prices are higher than 'ground', while the PDO certification is not eligible for 'ground' saffron. Beyond chemical organoleptic features (colour, flavour and aroma) there are quality controls than can only apply to saffron in strands, such as the length and brittleness of the filaments, as established by ISO standards (ISO 3632-1: 2011, revised in 2017) or the code of practice of 'La Mancha PDO'. Besides, ground saffron is more exposed to adulteration or fraud than saffron in strands. The attributes chosen in our experiment are presented in Table 2 and an example of a choice card is presented in Figure S1 in the online Appendix.

In the stated choice experiment, a third option was also included, phrased as 'What I bought the last time'. The inclusion of the third option is equivalent to an individual specific Status Quo (SQ), and this is particularly relevant in the context of saffron, where its substitution by a colourant (based on curcuma or other food colour additives) for culinary uses is frequent. A subsequent question inquired about the product bought by the respondent, saffron or colourant. If the product purchased the last time was saffron, the consumer indicated the brand name, from which, the origin, PDO or Organic labels, could be inferred. If colourant had been bought the last time, then all the attributes levels were null (no domestic production and absence of quality labels). In both cases, saffron and colourant, the respondents faced difficulties in remembering the price of the last purchase. For this reason, an average price for both saffron and colourant was assigned based on main retailers' online shopping websites.

In the Real Choice Experiment (RCE), it was not feasible to have a Status Quo option. Since this option is unique and specific for each respondent, it was not possible to provide a full range of products that consumers might have purchased the last time, and a two-option choice was used. Nevertheless, to address the possible bias induced by this 'forced choice' in the RCE, we compare the RCE results in the full sample (Sample 1) with that restricted to choice situations where the Status Quo option was not selected in the SCE (Sample 2). The logic is that, if the respondent had been given the option of an SQ alternative in the RCE, and the SQ was chosen in the SCE, then it is more likely that he/she would also have chosen SQ in the RCE. We used a D-efficient Multinomial (MNL) Logit[2] with a fractional factorial design for main effects, that led to 15 scenarios or choice sets grouped into two blocks (of eight and seven choice sets). Ngene software was used for this purpose.

2.3. The econometric model

In the usual theoretical framework defined by the random utility model (McFadden, 1974) each individual n faces a choice among J alternatives in choice situation t, and he/she obtains utility ($U^n_{j,t}$) from choosing alternative j over those available. Following

[2]A MNL model based-design has been demonstrated in the literature to provide consistent estimates also for Random Parameters Logit (RPL) estimation (Bliemer and Rose, 2010).

Table 2
Attributes and levels of the stated and real choice experiments

Attribute	Alternatives		Variable
Format	Strands (STR) Ground (GRO)		STR = 1 if the alternative is saffron in strands; = 0 otherwise GRO = 1 if the alternative is ground saffron; = 0 otherwise
Price (€/0.5 g)	1.5, 2.25 and 3€	1, 1.75 and 2.5€	PRICE
Origin	Teruel (Local) La Mancha (National) Iran (Imported)		TER = 1 if Origin is Teruel; = 0 otherwise LAM = 1 if Origin is La Mancha; = 0 otherwise
Quality label	Yes/No	No	PDO = 1 if carrying the PDO; = 0 otherwise
Organic certification	Yes/No	Yes/No	ORG = 1 if carrying Organic certification; = 0 otherwise

Note: The stated choice experiment included a third alternative Status Quo (SQ), specific for each respondent.

Lancaster (1966), the utility provided by a good can be decomposed into the utilities provided by its characteristics or attributes. These characteristics are observed while there are some other components of the individual's utility unobserved and are treated as random ($\varepsilon^n_{j,t}$). In our application, the utility is specified as:

$$U^n_{(j,t)} = \beta_{(0,j)} + \beta_1 \text{PRICE}^n_{(j,t)} + \beta_2 \text{TER}^n_{(j,t)} + \beta_3 \text{LAM}^n_{(j,t)} + \beta_4 \text{ORG}^n_{(j,t)} + \beta_5 \text{PDO}^n_{(j,t)} + \varepsilon^n_{(j,t)}. \quad (1)$$

In the SCE there are three alternatives (j = STR, GRO and SQ) while in the RCE, there are two (j = STR, GRO); $\beta_{0,j}$ is the alternative specific constant, that can only be identified for two alternatives in SCE (e.g. STR and GRO) and one in RCE (e.g. STR). In other words, the specific constant for one of the alternatives needs to be restricted to zero for identification. A constant is advisable for each labelled alternative to capture the effects of the label itself beyond those captured by the attributes (Lancsar *et al.*, 2017). More specifically, given the same characteristics in terms of price, origin or quality label, the utility of saffron in strands may differ from that of ground saffron. β_i (i = 1, …, 5) are coefficients to estimate or taste parameters; PRICE, TER, LAM, ORG and PDO are the observed characteristics as defined in Table 2. The utility for the Status Quo option in SCE is also explained by those same observed characteristics. This point is an important difference with respect to the modelling of the non-choice or opt-out option, where zero-level attributes are assumed (Hensher *et al.*, 2015, p. 53).

Assuming that the random term is *iid* (independently and identically) type I extreme value distributed, we obtain the conditional logit model. Allowing the taste parameters to vary across respondents (i.e. heterogeneous preferences), we get the random parameters mixed logit (RPL). In the RPL model, the utility function in (1) is modified by including the super-index n in the coefficients that are assumed to vary

randomly. The RPL model estimates the mean and standard deviation of the coefficients by maximum simulated likelihood (Train, 2003).

In our specification,[3] we assume that the parameters for TER, LAM, ORG and PDO follow a normal distribution in order to allow for preferences of opposite signs. The coefficient on price is kept fixed (i.e. non-random) in order to have finite variances in the marginal Willingness to Pay (WTP) measures, which in turn leads, in general, to more plausible WTP values (Hensher *et al*., 2015). The marginal WTP is calculated as the ratio between the characteristic and the price coefficients (with sign changed), and indicates how much the individual is willing to pay in order to gain some characteristic, while keeping the utility constant (Hensher *et al*., 2015, p. 379). With the fixed price assumption, we are ruling out that different consumers can have different preferences towards price. Thus, keeping the price fixed may result in a loss of realism in favour of a gain of modelling convenience (Hole and Kolstad, 2012) that explains why it is a commonly used (e.g. Lusk, 2003; Tonsor and Shupp, 2011).[4]

2.4. Willingness to Pay: Ccomparison across experimental conditions

Using mean coefficient estimates from the stated and real experiments, we calculate the unconditional population mean, what gives a first indication of the dependence of WTP on the modelling approach, and of any hypothetical bias. A second, richer approach estimates taste coefficients for each individual within the sample conditional on the pattern of choices observed for each respondent (see Train, 2003, Chapter 11 for details on the procedure).[5] The behavioural implications of these individual conditional parameters (impossible with unconditional estimates) can be used for segmentation purposes (Hensher *et al*., 2015, p. 650). More specifically, dividing the individual taste parameter by the (fixed) price coefficient, we calculate the Marginal WTP for the characteristics of interest (TER, LAM, ORG, PDO) for each individual in the sample. Of particular interest are the relationships between individual WTP across experimental conditions and individuals' characteristics, which can inform the existence and factors influencing hypothetical bias. For this purpose, we conduct the following bivariate analysis:

1 Pairwise correlation between stated and real WTP for each of the characteristics, to test for respondents' consistency, to check if higher stated WTP for one attribute in the stated choice experiment is also accompanied by higher WTP in the non-hypothetical setting. Under non-normality, Spearman's rank correlation is used, and normality is assessed with the Shapiro–Wilks test.

---

[3]A more flexible specification where the random taste parameters are allowed to be correlated (RPL-C) was also checked (Revelt and Train, 1998). Allowing for correlation improved the explanatory power slightly in any experimental setting. However, the RPL-C models, either unrestricted or restricted, led to relatively implausible WTP distributions. Accordingly, we decided to sacrifice some modelling flexibility in favour of economic plausibility. The RPL-C model results are available from the authors upon request.

[4]The choice model in WTP space (Scarpa *et al*., 2008) would solve this restriction, but empirical results are not unanimous on its superiority in terms of plausibility of WTP distributions while in most cases, the models in preference space tend to fit the data better (see Hole and Kolstad, 2012).

[5]All estimations are conducted in Stata 15.0 using commands developed by Hole (2007): *mixlogit* for model estimation and *mixlbeta* to obtain individual-level coefficients.

2 A test for differences in WTP elicited in the stated and real experiments to identify hypothetical bias. Under the absence of normality, the non-parametric Wilcoxon matched-pairs signed-ranks test is applied, that tests the null of equal medians.
3 A test for differences in WTP elicited in the stated and real choice experiments for each specific attribute of interest (TER, ORG and PDO) across individuals segmented according to socio-demographics or product-related knowledge and experience variables, using the non-parametric Kruskal–Wallis test.

3. Results and Discussion

3.1. Saffron knowledge, use and choices

Table 3 presents the relative frequencies for the selected variables. Comparing our results with those of Sánchez *et al*. (2018), the pattern of consumption and purchase habits has remained rather stable: the ratio of consumers (CONS) is 75.5% (vs. 78% in their study); the percentage of occasional consumers is identical, 52%; saffron is mainly used as a condiment and in rice dishes; and is bought mostly at super and hypermarkets (78% vs. 74% in their study).

Only 23% of respondents consume saffron on a weekly or monthly regular basis (HFREQ). Interestingly, most of the respondents know that saffron is produced in the same region (78%) and accurately locate the area (63%) (KNOWTer), while only a minority is aware of the production in other regions (36%) (KNOWSp), La Mancha in particular (25%), and only a minority knows about the PDO La Mancha saffron (13%). In terms of certification labels available for food products, the organic label (KNOWOrg) is more widely known than the PDO/PGI (59% vs. 46%) (KNOWPDO). The low salience of PDO, however, increases significantly when mentioned in conjunction with particular products, which explains why the percentage of respondents that have ever bought a PDO (or Organic) food product (PURCHPDO/Org) rises to 83%.

In the SCE, saffron in 'strands' was chosen in 54% of the occasions, 'ground' in 31% and the Status Quo in 14% of the choice situations. When the Status Quo was chosen, colourant had been bought in 52% of the choice situations and saffron in the remaining 48%. In addition, within the choice of Status Quo, when saffron had been purchased, this was always conventional or non-organic; in 19% of the choices, saffron was local (from Teruel), in 17% saffron was from La Mancha and in 3% of the choices La Mancha PDO had been bought. Likewise, 37.5% of the respondents chose the Status Quo at least once, and from these, 86% recalled the product (saffron brand included) of the last purchase. Only 11 consumers who reported choosing saffron as their last purchase did not recall the brand and missing values for all products characteristics were assigned.[6] There are only five respondents who chose the SQ in every choice situation and these were eliminated from the analysis in both experiments.

[6]We failed to find a statistical association between choosing the status quo at least once and the frequency of consumption; or between the frequency of consumption and the ability to remember the last purchased product. In other words, more regular consumers of the product do not tend to choose the last product bought more often; and more regular consumers do not tend to remember the last product bought significantly better. Accordingly, within our sample of consumers, we do not consider that there is a serious problem of endogeneity, as suggested by one referee.

Table 3
Variables definition and percentage in the sample

Variable	Description	%
FEMALE	1 = female; 0 = male	63.9
OLD	1 = if ≥45 years old; 0 = otherwise	51.0
MARRIED	1 = Living with a partner, with or without children; 0 = otherwise	66.3
UNIV	1 = University degree; 0 = otherwise	45.7
LINC	1 = low household income (<1,500€); 0 = otherwise	29.8
ROOTS	1 = Time living in Aragón, all or most part of his/her life; 0 = otherwise	91.8
CONS	1 = consumer of saffron; 0 = otherwise (never)	75.5
HFREQ	1 = consumption of saffron at least once a month; 0 = otherwise	22.6
KNOWTer	1 = Knowledge about the cultivation of saffron in Teruel; 0 = otherwise	63.5
KNOWSp	1 = Knowledge about the cultivation of saffron in other Spanish regions (≠ Teruel); 0 = otherwise	36.5
KNOWPDO	1 = Knowledge of PDO/PGI certifications; 0 = otherwise	46.1
KNOWOrg	1 = Knowledge of Organic certification; 0 = otherwise	58.6
PURCHPDO/Org	1 = Purchase, at least once, of a certified food with PDO/PGI or Organic	82.7

3.2. Estimation results

For comparison purposes, we also used a basic conditional logit (CL). Results for the CL and RPL models are presented in Table 4. All specifications in both experimental conditions and samples result in consistent patterns: utility increases with lower prices (i.e. normal demand); saffron in 'strands' is preferred over 'ground'; the local origin TER is preferred over the national origin LAM, and these over the imported variety; an Organic/PDO certification is preferred over the lack of such certifications; and the marginal utility of PDO is higher than that of ORG. Furthermore, the (population) WTP gap between stated and real experiments, for each model specification, also move in the same direction; positive (i.e. stated WTP> real WTP) for all characteristics except for STR.[7]

The Pseudo-R^2 indicates that the most flexible model RPL fits the data better than CL, and the BIC (Bayesian Information Criterion) is also minimised in the RPL model, in all sample and experimental conditions. Likewise, the Wald test favours the random over the fixed parameters specification. Significant standard deviations are found for ORG and PDO characteristics. Interestingly, preferences for the local and national origins (TER and LAM) are heterogeneous only in the stated experiment, implying that preferences become more homogeneous in the non-hypothetical setting, which is consistent with Grebitus *et al.* (2013).

[7]Comparison between STR in RCE with (STR-GRO) in SCE.

Table 4
Estimation results: Stated and real choice experiments

	SCE		RCE: Sample 1		RCE: Sample 2	
	CL	RPL	CL	RPL	CL	RPL
Mean						
STR	0.86 (0.00)	1.17 (0.00)	0.64 (0.000)	0.81 (0.00)	0.80 (0.00)	1.02 (0.00)
GRO	0.40 (0.03)	0.55 (0.02)	—	—		
PRICE	−0.47 (0.00)	−0.67 (0.00)	−0.56(0.001)	−0.78 (0.00)	−0.71 (0.00)	−0.99 (0.00)
TER	1.16 (0.00)	1.60 (0.00)	1.06 (0.000)	1.48 (0.00)	0.93 (0.00)	1.34 (0.00)
LAM	0.77 (0.00)	0.88 (0.00)	0.52 (0.009)	0.50 (0.04)	0.42 (0.06)	0.38 (0.15)
ORG	0.77 (0.00)	0.92 (0.00)	0.35 (0.009)	0.37 (0.02)	0.31 (0.03)	0.31 (0.08)
PDO	0.94 (0.00)	1.50 (0.24)	0.75 (0.000)	1.29 (0.00)	0.65 (0.00)	1.18 (0.00)
SD						
TER	—	1.46 (0.00)	—	1.33 (0.11)		1.32 (0.19)
LAM	—	0.79 (0.00)	—	0.27 (0.53)		0.35 (0.48)
ORG	—	0.79 (0.00)	—	0.54 (0.00)		0.59 (0.00)
PDO	—	1.68 (0.00)	—	1.56 (0.00)		1.62 (0.00)
LL	−1,135.36	−1,092.50	−823.42	−809.69	−698.01	−685.72
Wald χ^2	359.46 (0.00)	258.35 (0.00)	196.05 (0.00)	151.61 (0.00)	188.32 (0.00)	109.15 (0.00)
BIC	2,329.38	2,277.18	1,694.96	1,699.61	1,443.36	1,450.33
R^2	0.28	0.31	0.22	0.23	0.25	0.26
No. obs. (N.respondents)	4,355 (203)	4,355 (203)	3,048 (203)	3,048 (203)	2,670 (203)	2,670 (203)
Population WTP (mean = median) (€/0.50 g)						
STR	1.81 (0.00)	1.75 (0.00)	1.13 (0.00)	1.04 (0.00)	1.12 (0.00)	1.03 (0.00)
GRO	0.83 (0.03)	0.83 (0.01)	—	—		
TER	2.45 (0.00)	2.40 (0.00)	1.88 (0.05)	1.90 (0.03)	1.31 (0.06)	1.36 (0.03)
LAM	1.62 (0.01)	1.31 (0.01)	0.93 (0.12)	0.64 (0.15)	0.58 (0.18)	0.38 (0.26)
ORG	1.63 (0.01)	1.38 (0.00)	0.63 (0.12)	0.47 (0.13)	0.43 (0.15)	0.32 (0.18)
PDO	1.98 (0.00)	2.24 (0.00)	1.33 (0.02)	1.65 (0.01)	0.91 (0.02)	1.19 (0.01)
Relative importance						

Table 4
(*Continued*)

	SCE		RCE: Sample 1		RCE: Sample 2	
	CL	RPL	CL	RPL	CL	RPL
Format	11%	14%	15%	19%	19%	24%
Price	22%	31%	26%	36%	33%	46%
Origin	27%	37%	25%	35%	22%	31%
Organic	18%	22%	8%	9%	7%	7%
Quality label	22%	35%	18%	30%	15%	28%

Notes: *P*-value in parentheses, based on robust standard errors clustered per individual. Wald χ^2 test for random against a fixed parameters model (*P*-value in parentheses). McFadden Pseudo $R^2 = 1 - (LL/L0)$ where L0 is the log-likelihood in the intercepts-only model. STR and GRO alternative specific constants in SCE need to be interpreted as relative to the (excluded) SQ constant; in RCE, the alternative specific constant on STR is relative to GRO. Relative Importance for each attribute is calculated as the ratio between the range of marginal utilities for each attribute (i.e. max − min) and the sum of ranges for all attributes.

Restricting the sample (from Sample 1 to Sample 2) in RCE does not have any impact on the rankings of attributes according to their relative importance[8] or population mean WTP for characteristics. Thus, price ranks first, followed by origin, quality label, format and production technique. From the highest to the lowest population mean WTP, the characteristics are ordered as: TER, PDO, STR, LAM and ORG. Besides, we do not find any significant correlation between the number of times the consumer chooses the SQ in SCE and the individual WTP elicited in the real experiment. Accordingly, subsequent analysis of real WTP will be based on estimation results from Sample 1.[9]

Within the SCE, the ranking agrees with RCE in placing TER and PDO with the first and second highest WTP, while the remaining characteristics, from highest to lowest WTP are ordered slightly differently: ORG, LAM and STR. It is interesting to note that the 'strands' presentation gains salience in the real choice: the relative importance of the format goes up from 14% in the stated experiment to 19–24% in real choice, and (population) WTP also rises from €0.92 to €1.04–1.03. Similarly, price becomes more relevant in the real experiment (36–46% vs. 31% in the stated experiment), consistent with Grebitus *et al.* (2013). On the other hand, the Organic certification loses importance when comparing preferences in the stated (22%) and real experimental conditions (9–7%). The Quality Label, on the other hand, as well as Origin, remains quite stable across experimental conditions, confirming a trend observed in the marketing literature for greater preference towards local versus organic food (e.g. Bazzani *et al.*, 2017).

Based on the parameters estimates of the RPL model in Table 4 (Sample 1) we simulate individual WTP conditional on the pattern of the respondent's selection. The descriptive statistics of WTP for each of the characteristics are shown in Table 5. The conditional estimation provides in general non-normal distributed WTP, in particular in the stated choice experiment, and consequently, the median may be a more adequate central tendency measure. Furthermore, the median has been considered a more useful measure than the mean from a societal or a market research perspective (Balcombe *et al.*, 2009), as it implies that the majority of the consumers agree to pay this specific amount of money. Accordingly, in Table 5 the median is also included.

Confirming the trend observed with the population mean WTP (Table 4), and according to the Wilcoxon signed-ranks test, the conditional WTP also reveals a significantly higher WTP in the stated choice experiment than in the real experiment for

[8]Relative Importance for each attribute (e.g. origin) is calculated as the ratio between the range of marginal utilities for each attribute (i.e. max − min, e.g. max in origin corresponds to TER, and minimum to the absent level, IRN, and equal to 0) and the sum of ranges for all attributes (Maaya *et al.*, 2018, p. 11).

[9]There is lack of sufficient literature dealing with the effect of including or excluding an opt-out on a forced choice in the choice experiment (Penn *et al.*, 2019). Moreover, the type of products evaluated (private or public goods, etc.) and the choice design are not uniform, making the results difficult to compare. In our paper, the forced choice does not appear to have a significant effect on the main results. These results are, at least to a certain extent, in line with Carlsson *et al.* (2007), who evaluating the impact of animal welfare improvements (credence attribute) on consumer's meat (beef and chicken) choices found no significant differences between the forced and unforced models in WTP, in contrast with Kallas and Gil (2012) who investigated consumer choices toward differentiated rabbit meat. Nevertheless, further research investigating the impact of a forced choice on credence or labelling agro-food attributes is needed.

Table 5
Conditional WTP in stated (SCE) and real choice experiment (RCE) (€/0.50 g)

	SCE					RCE					SCE–RCE
	Mean	Med.	SD	p5	p95	Mean	Med.	SD	p5	p95	Corr.
STR vs. GRO	0.92	0.92	0.00	0.92	0.92	1.04	1.04	0.00	1.04	1.04	—
TER***	2.42	2.25	1.27	0.06	4.09	1.89	1.63	0.83	0.68	2.84	0.24 (0.00)
LAM***	1.34	1.37	0.61	0.30	2.15	0.64	0.65	0.09	0.48	0.75	0.43 (0.00)
ORG***	1.39	1.33	0.65	0.39	2.52	0.47	0.48	0.32	0.01	1.08	0.40 (0.00)
PDO***	2.27	2.70	1.53	−0.69	4.13	1.69	2.17	1.12	−0.60	2.94	0.45 (0.00)

Notes: Med: median; SD: standard deviation; p5: percentile 0.05; p95: percentile 95; Corr: Spearman correlation between conditional WTP in stated and real experiments (*P*-value in parentheses).
***Indicates significant median differences at 1% for WTP between stated and real experiments using the non-parametric Wilcoxon signed-ranks test

all characteristics except for the format in strands. Our ratios of WTPs in the stated versus the real experiments are in general below those reported in the meta-analysis reviewed by Loomis (2011) (around 3), with the exception of the organic label (2.77). The local TER (1.38) and PDO (1.24) are however more aligned to specific choice experiments applied to food. For instance, Lusk and Schroeder (2004) report an average of 1.2; De Magistris *et al*. (2013) around 1.5. Yue and Tong (2009) report a hypothetical bias of around 7.5% and 9% (i.e. ratios of 1.075 and 1.09) for organic and local attributes, respectively, in fresh tomatoes. Interestingly though, these authors displayed the real product in both experiments although the transaction only took place in the non-hypothetical experiment. The authors argue that the display of the product helps to reduce the hypothetical bias in comparison with a stated experiment where only pictures are presented, which is in accordance with their lower ratio in comparison with the above-mentioned results and our own estimates.

Comparing across characteristics, the PDO enjoys the highest (median) WTP in any experimental condition, followed by the local origin TER. The WTP values for the national origin LAM and the organic certification are much lower in any condition. For instance, focusing on the results provided by the real choice, the median WTP for PDO is €2.17 and TER €1.63, while for LAM, WTP falls to €0.65 and ORG €0.48. Consequently, it is clear that the local speciality of saffron has an important appeal amongst consumers that could be reinforced by certification with a PDO. Saffron is a product hardly investigated from the perspective of consumers' preferences (with the exceptions of Sanjuán-López *et al*., 2009, 2010). Nevertheless, this result is consistent with previous results where the quality label is found to act as a powerful quality cue, especially when attached to a recognised origin (Resano *et al*., 2012). The organic label, on the other hand, evokes a WTP of much lower magnitude. In this sense, our results are in line with the strand of the literature on sustainable food consumption that finds evidence of the superior value attached to the local over the organic claim (Denver and Jensen, 2014; Gracia *et al*., 2014; De Magistris and Gracia, 2016).

It is also worth noting that the dispersion of WTP is narrower in the real experiment (lower standard deviations and narrower range between the lower and upper 5 percentiles). This implies that a more realistic budget constraint accompanied with the

display of the product leads to a lower variability of WTP across consumers, revealing relatively closer preferences. In our case, display of the real product especially favours the choice of saffron in strands versus ground and the consequent WTP for this characteristic is higher in the real than in the hypothetical experiment. Sanjuán-López *et al.* (2009) found that the format in strands implies a higher implicit price than powder using a hedonic approach.

All the Spearman correlations between conditional WTP for individual attributes between both choice experiments are positive and significant although not of a high magnitude (last column in Table 5). Therefore, respondents show a certain degree of consistency, i.e. consumers with higher WTP in the stated or hypothetical condition also show a higher WTP in the non-hypothetical choice condition. Interestingly, the correlation is lower for the local speciality (0.24) and the organic certification (0.40), and greatest (0.45) for PDO. This implies that the local and organic attributes have a resonance amongst consumers, in general, that does not translate into a more committed/involved choice for all of them. This could be explained by the 'social desirability' bias, or the tendency to answer according to what is perceived as a socially desirable in order to project a favourable self-image.

Finally, we tried to investigate some of the underlying personal factors that may contribute to explaining the (conditional) WTP gap between the stated and real experiments. Empirical evidence on this gap is scarce, although Grebitus *et al.* (2013) found that certain psychographic characteristics influence hypothetical bias, such as agency and openness. Interestingly, self-confident or dominant individuals tended to reveal greater hypothetical bias, while creative or open consumers showed a smaller bias. Moreover, Hofstetter *et al.* (2013) demonstrated that consumers' ability and motivation factors may influence the hypothetical bias for innovative products. In particular, and contrary to expectations, consumers with previous product category knowledge may have a tendency to overestimate their WTP in a hypothetical context. Here, we are only able to explore the set of socio-demographic and product-related variables described in Table 3, while only significant results (based on the Kruskal–Wallis test) are highlighted in Table 6. *A priori*, we would expect that certain socio-demographic variables, such age or education, as well as familiarity with the product (CONS and HFREQ) might help to mitigate the gap, while certain knowledge and previous experience with labels (KNOWPDO, KNOWOrg) might help to reduce the bias for the PDO and ORG characteristics. However, we failed to find a profile of respondents that consistently tend to overstate their WTP. For instance, we find that being older ($P < 0.05$) or having spent most of one's life in the region ($P < 0.10$) is associated with a bigger gap in WTP for the local variety (TER), increasing the gap by €0.06 and €0.35, respectively; the WTP gap for Organic certification is lower for married ($P < 0.10$) and higher-income respondents ($P < 0.05$), as well as for those knowing about the cultivation of saffron in the region ($P < 0.01$) (€0.15, €0.13, and €0.31, respectively); while being married ($P < 0.05$) or knowing about the local variety ($P < 0.10$) increases the WTP gap for the PDO certification (€0.18 and €0.20, respectively).

Our findings suggest that consumers relatively more knowledgeable and with deeper roots in the territory tend to show a higher hypothetical bias on the region of origin and its certification, probably driven by affective and normative food dimensions, more intense in a hypothetical context. Thus, the own region of origin (van Ittersum *et al.*, 2003) and the PDO (Resano *et al.*, 2012) may affect consumers' choices not only as a quality indicator, but also through the sense of belonging to the region, or as a

Table 6
Conditional WTP gap (stated − real) per subgroup of consumers (€/0.50 g)

	Stated-Real WTP gap		
Variable	TER	ORG	PDO
OLD	**0.84 (0.78) ****	0.94 (0.92)	0.82 (0.71)
MARRIED	0.79 (0.88)	**0.86 (1.01) ***	**0.82 (0.64)****
LINC	0.88 (0.76)	**1.01 (0.88)****	0.71 (0.80)
ROOTS	**0.85 (0.50)***	0.93 (0.98)	0.78 (0.66)
KNOWTER	0.84 (0.79)	**0.82 (1.13)*****	**0.83 (0.63)***
Median	0.83	0.93	0.77

Notes: Outside parentheses, the gap of median WTP when 'Variable' = 1; in parentheses, the gap of median WTP when 'Variable' = 0
***, ** and *denote significant differences at 1%, 5% and 10%, respectively, according to the Kruskal–Wallis non-parametric test, across subgroups of consumers defined by 'Variable'.

way to promote its rural development, which could be especially relevant in the case of a 'less favoured area' such as Teruel. The impact of these components might be weaker in an organic food product, although further research is needed to corroborate this hypothesis.

4. Concluding Remarks

Saffron has been historically cultivated in the province of Teruel (Spain), but socio-demographic forces have substantially reduced production since the beginning of the 20th century. Saffron is well suited to the ecological and climatic conditions of the area, helping to preserve biodiversity. The extent to which it can also contribute to socio-economic sustainability, providing an alternative or complementary source of income to arrest the long decline in rural depopulation, depends on the existence of potential demand. In this paper, we have investigated whether certain specific attributes, such as the origin and labelling, as well as the organic feature, could help to promote the demand for local saffron amongst local consumers, and to identify the price premium that they would be willing to pay. Furthermore, we have investigated whether some consumers may behave differently in a hypothetical versus a real experiment and whether particular traits systematically explain this gap.

The international market is largely lead by price, where Iran is the indisputable leader. However, niche gourmet international markets could also be more attainable through verified claims that would better match the demand for authenticity and superior quality. Our results suggest that preferences move toward lower prices, local origin, quality certifications and the use of organic production techniques. According to our WTP estimates, regional producers should focus on the supply of saffron in strands and the explicit recognition and promotion of the local Teruel origin. Organic production, on the other hand, does not elicit a substantial price differential for saffron, providing further evidence of the superior value attached to the local rather than the organic claim. The pursuit of a PDO is also supported by the data but further analysis of the costs induced by the certification would be needed in order to provide a stronger recommendation. Besides, the relatively small level of production may make it difficult to achieve the PDO designation.

From a methodological perspective, eliciting preferences in a real experiment, where an economic incentive exists, the product is displayed and a real purchase takes place, has led in our study to significantly lower WTP and more homogeneous preferences, adding evidence to the existence of an upward hypothetical bias in choice experiments. Nevertheless, although there is a hypothetical bias in all characteristics (with the exception of the format), consumers tend to be consistent, i.e. consumers with higher WTP for a specific attribute in the hypothetical setting also manifest a higher WTP in the real experiment. Interestingly, this consistency is significantly lower and the hypothetical bias significantly higher for the local origin and the organic attributes, which might reflect social desirability bias. On the other hand, the strands presentation diminishes the possibility of getting a fraudulent product, and this characteristic seems to outweigh the convenience offered by the ground format, leading to a higher WTP for saffron in strands in comparison with ground saffron and also reversing the hypothetical bias (i.e. WTP for strands is higher in the real than in the stated experiment).

Findings also suggest that consumers relatively more knowledgeable and with deeper roots in the territory tend to overstate their WTP for the local product and its labelling, manifesting a greater hypothetical bias, probably caused by the stronger impact of the affective and normative dimensions inherent to food consumption. These results provide useful insights for practitioners, who need to decide on the use of stated or real experiments, and can be a starting point for further analysis of the influence of personal characteristics on the hypothetical bias. This could also be relevant when deciding the target population of the experiment or the sampling procedure. Furthermore, the use of an individual Status Quo alternative in the hypothetical experiment has proved to be useful in our case, characterising more accurately the market reality when close substitutes exist. Further investigation is needed to explore the virtues of the individual Status Quo in comparison with the more common 'non-choice' or 'opt-out' alternatives.

Finally, we should note that, although we have tried to address the possible bias induced by the forced choice in the Real Choice Experiment, when comparing WTP between experimental conditions, we cannot completely isolate the effect of this forced choice from the effect of hypothetical bias. Exploring the underlying motivation that induces bias,[10] as well as including a wider range of consumers' characteristics, in particular, psycho-graphic and personal traits, is a promising future extension of this research.

Supporting Information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Figure S1: Example of a choice set (Stated Choice Experiment).

References

Alfnes, F. and Steine, G. *None-of-these Bias in Hypothetical Choice Experiments*, Discussion Paper No. DP-06/05 (Department of Economics and Resources Management, Norwegian university of Life Sciences, Aas, 2005).

[10]We thank an anonymous referee for raising this point.

Alfnes, F., Guttormsen, A. G., Steine, G. and Kolstad, K. 'Consumers' willingness to pay for the color of salmon: A choice experiment with real economic incentives', *American Journal of Agricultural Economics*, Vol. 88, (2006) pp. 1050–1061.

Alvarez, J. M. and Mallor, C. 'Evaluación de parámetros de calidad del azafrán del Jiloca (Teruel)', *ITEA*, Vol. 105, (2009) pp. 61–66.

Ávila, G. J. *Historia del Azafrán* (Barcelona, Spain: Ediciones Zendrera Zariquiey, 1999).

Balcombe, K., Chalak, A. and Fraser, I. 'Model selection for the mixed logit with Bayesian estimation', *Journal of Environmental Economics and Management*, Vol. 57, (2009) pp. 226–237.

Bathaie, S. Z. and Mousavi, Z. 'New applications and mechanisms of action of saffron and its important ingredients', *Critical Reviews in Food Science and Nutrition*, Vol. 50, (2010) pp. 761–786.

Bazzani, C., Caputo, V., Nayga, R. and Canavari, M. 'Revisiting consumer's valuation for local versus organic food using a non-hypothetical choice experiment: Does personality matter?', *Food Quality and Preference*, Vol. 62, (2017) pp. 144–154.

Bernués, A., Olaizola, A. and Corcoran, K. 'Extrinsic attributes of red meat as indicators of quality in Europe: An application for market segmentation', *Food Quality and Preference*, Vol. 14, (2003) pp. 265–276.

Bliemer, M. C. J. and Rose, J. M. 'Construction of experimental designs for mixed logit models allowing for correlation across choice observations', *Transportation Research Part B: Methodological*, Vol. 44, (2010) pp. 720–734.

Carlsson, F., Frykblom, P. and Lagerkvist, C. J. 'Consumer willingness to pay for farm animal welfare: Mobile abattoirs versus transportation to slaughter', *European Review of Agricultural Economics*, Vol. 34, (2007) pp. 321–344.

Charness, G., Gneezy, U. and Kuhn, M. A. 'Experimental methods: Between-subject and within-subject design', *Journal of Economic Behavior & Organization*, Vol. 81, (2012) pp. 1–8.

De Magistris, T. and Gracia, A. 'Consumers' willingness-to-pay for sustainable food products: The case of organically and locally grown almonds in Spain', *Journal of Cleaner Production*, Vol. 118, (2016) pp. 97–104.

De Magistris, T., Gracia, A. and Nayga, R. M. 'On the use of honesty priming tasks to mitigate hypothetical bias in choice experiments', *American Journal of Agricultural Economics*, Vol. 95, (2013) pp. 1136–1154.

Denver, S. and Jensen, J. D. 'Consumer preferences for organically and locally produced apples', *Food Quality and Preference*, Vol. 31, (2014) pp. 129–134.

Gracia, A. 'Consumers' preferences for a local food product: a real choice experiment', *Empirical Economics*, Vol. 47, (2014) pp. 111–128.

Gracia, A., Barreiro-Hurlé, J. and López-Galán, B. 'Are local and organic claims complements or substitutes? A consumer preferences study for eggs', *Journal of Agricultural Economics*, Vol. 65, (2014) pp. 49–67.

Grebitus, C., Lusk, J. L. and Nayga, R. M. 'Explaining differences in real and hypothetical experimental auctions and choice experiments with personality', *Journal of Economic Psychology*, Vol. 36, (2013) pp. 11–26.

Hensher, D. A., Rose, J. M. and Greene, W. H. *Applied Choice Analysis*, 2nd edn (Cambridge, UK: Cambridge University Press, 2015).

Hofstetter, R., Miller, K. M., Krohmer, H. and Zhang, Z. J. 'How do consumer characteristics affect the bias in measuring willingness to pay for innovative products?', *The Journal of Product Innovation Management*, Vol. 30, (2013) pp. 1042–1053.

Hole, A. R. 'Fitting mixed logit models by using maximum simulated likelihood', *Stata Journal*, Vol. 7, (2007) pp. 388–401.

Hole, A. R. and Kolstad, J. 'Mixed logit estimation of willingness to pay distributions: A comparison of models in preference and WTP space using data from a health-related choice experiment', *Empirical Economics*, Vol. 42, (2012) pp. 445–469.

IAEST. *Padrón municipal de habitantes* (Zaragoza: Instituto Aragonés de Estadística (IAEST), 2010a). Available at: http://servicios3.aragon.es/iaeaxi/tabla.do?path=/02/01/01/01/&file=0201010101.px&type=pcaxis&L=0 (last accessed 5 June 2019).

IAEST. *Datos Básicos de Aragón* (Zaragoza: Instituto Aragonés de Estadística (IAEST), 2010b). Available at: http://www.aragon.es/DepartamentosOrganismosPublicos/Institutos/InstitutoAragonesEstadistica/AreasGenericas/DatosBasicos/ci.02_Historicos.detalleDepartamento?channelSelected=0 (last accessed 5 June 2019).

IAEST. *Datos Básicos de Aragón* (Zaragoza: Instituto Aragonés de Estadística (IAEST), 2011). Available at: http://www.aragon.es/DepartamentosOrganismosPublicos/Institutos/InstitutoAragonesEstadistica/AreasGenericas/DatosBasicos/ci.02_Historicos.detalleDepartamento?channelSelected=0 (last accessed 5 June 2019).

van Ittersum, K., Cander, M. J. J. M. and Meulenberg, M. T. G. 'The influence of the image of a product's region of origin on product evaluation', *Journal of Business Research*, Vol. 56, (2003) pp. 215–226.

Kallas, Z. and Gil, J. M. 'A dual response choice experiments (DRCE) design to assess rabbit meat preferences in Catalonia: A heteroscedastic extreme-value model', *British Food Journal*, Vol. 114, (2012) pp. 1394–1413.

Kyriakoudi, A., Ordoudi, S. A., Roldán-Medina, M. and Tximidou, M. Z. 'Saffron, a functional spice', *Austin Journal of Nutrition and Food Sciences*, Vol. 3, (2015) pp. 1059.

Lancaster, K. J. 'A new approach to consumer theory', *Journal of Political Economy*, Vol. 74, (1966) pp. 132–156.

Lancsar, E., Fiebig, D. G. and Hole, A. R. 'Discrete choice experiments: A guide to model specification, estimation and software', *Pharmacoeconomics*, Vol. 35, (2017) pp. 697–716.

Loomis, J. 'What's to know about hypothetical bias in stated preference valuation studies?', *Journal of Economic Surveys*, Vol. 25, (2011) pp. 363–370.

Louviere, J., Hensher, D. and Swait, J. *Stated Choice Methods: Analysis and Application* (Cambridge, UK: Cambridge University Press, 2000).

Lusk, J. L. 'Effects of cheap talk on consumer willingness-to-pay for golden rice', *American Journal of Agricultural Economics*, Vol. 85, (2003) pp. 840–856.

Lusk, J. L. and Schroeder, T. C. 'Are choice experiments incentive compatible? A test with quality differentiated beef steaks', *American Journal of Agricultural Economics*, Vol. 86, (2004) pp. 467–482.

Maaya, L., Meulders, M., Surmont, N. and Vandebroek, M. 'Effect of environmental and altruistic attitudes on willingness-to-pay for organic and fair trade coffee in Flanders', *Sustainability*, Vol. 10, (2018) pp. 1–21.

MARM. Anuario Estadístico 2010 (Madrid, Spain: Subdirección General de Estadística, 2010). Available at: http://www.marm.es/es/estadistica/temas/anuario-de-estadistica/ (last accessed 5 June 2019).

MARM. Anuario Estadístico 2017 (Madrid, Spain: Subdirección General de Estadística, 2017). Available at https://www.mapa.gob.es/es/estadistica/temas/publicaciones/anuario-de-estadistica/2017/default.aspx?parte=3&capitulo=13&grupo=4&seccion=16 (last accessed 5 June 2019).

Marsden, T. and Smith, E. 'Ecological entrepreneurship: Sustainable development in local communities through quality food production and local branding', *Geoforum*, Vol. 36, (2005) pp. 440–451.

McFadden, D. 'Conditional logit analysis of qualitative choice behavior', in P. Zarembka (ed.), *Frontiers in Econometrics* (New York: Academic Press, pp. 105–142, 1974).

Olesen, I., Alfnes, F., Røra, M. B. and Kolstad, K. 'Eliciting consumers' willingness to pay for organic and welfare-labelled salmon in a non-hypothetical choice experiment', *Livestock Science*, Vol. 127, (2010) pp. 218–226.

PDR. *Spain – Rural Development Programme (Regional) – Aragon 2014–2020. The European Agricultural Fund for Rural Development, Europe Investing in Rural Areas* (Aragon: Gobierno de Aragón, Departamento de Agricultura, Ganadería y Medio Ambiente, 2016).

Pedersen, L. B. and Gyrd-Hansen, D. 'The use of status quo and opt out options in choice experiments. Implications of researchers' dubious use of the 'neither' option', International Choice Modelling Conference, May 2013.

Penn, J. M., Hu, W. and Cox, L. J. 'The effect of forced choice with constant choice experiment complexity', *Journal of Agricultural and Resource Economics*, Vol. 44, (2019) pp. 439–455.

Peschel, A. O., Grebitus, C., Alemu, M. A. and Hughner, R. S. 'Personality traits and preferences for production method labeling – A latent class approach', *Food Quality and Preference*, Vol. 74, (2019) pp. 163–171.

Resano, H., Sanjuán, A. I. and Albisu, L. M. 'Consumers' response to the EU quality policy allowing for heterogeneous preferences', *Food Policy*, Vol. 37, (2012) pp. 355–365.

Resano, H., Olaizola, A. M. and Dominguez-Torreiro, M. 'Exploring the influence of consumer characteristics on veal credence and experience guarantee purchasing motivators', *Meat Science*, Vol. 141, (2018) pp. 1–8.

Revelt, D. and Train, K. 'Mixed logit with repeated choices: Households' choices of appliance efficiency level', *Review of Economics and Statistics*, Vol. 80, (1998) pp. 647–657.

Sánchez, A. M., Gracia, A., Mallor, C. and Jurado, F. *Rural Economics in the Spanish Region of Aragon. Local Intention to use High Quality Saffron in Cookery*. XII Congreso Iberoamericano de Estudios Rurales, Segovia, July 2018.

Sanjuán-López, A., Resano-Ezcaray, H. and Camarena-Gómez, D. 'Developing marketing strategies for Jiloca saffron: A price hedonic model', *Spanish Journal of Agricultural Research*, Vol. 7, (2009) pp. 305–314.

Sanjuán-López, A. I., Philippidis, G. and Resano-Ezcaray, H. 'How useful is acceptability to explain economic value? An application on the introduction of innovative saffron products into commercial markets', *Food Quality and Preference*, Vol. 22, (2011) pp. 255–263.

Sanjuán-Lopez, A. I., Zeballos, G. and Resano Ezcaray, H. *El uso actual y potencial del azafrán entre los consumidores aragoneses* (Zaragoza: CITA Vol. 54, 2010). Available at: https://citarea.cita-aragon.es/citarea/bitstream/10532/1502/1/10532-80_1.pdf (last accessed 5 June 2019).

Scarpa, R., Thiene, M. and Train, K. 'Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps', *American Journal of Agricultural Economics*, Vol. 90, (2008) pp. 994–1010.

Silva, A., Nayga, R., Benjamin, L., Campbell, B. L. and Park, J. L. 'Revisiting cheap talk with new evidence from a field experiment', *Journal of Agricultural and Resource Economics*, Vol. 36, (2011) pp. 280–291.

Toler, S., Briggeman, B. C., Lusk, J. L. and Adams, D. C. 'Fairness, farmers markets, and local production', *American Journal of Agricultural Economics*, Vol. 91, (2009) pp. 1272–1278.

Tonsor, G. T. and Shupp, R. S. 'Cheap talk scripts and online choice experiments: "Looking beyond the mean"', *American Journal of Agricultural Economics*, Vol. 93, (2011) pp. 1015–1031.

Train, K. E. *Discrete Choice Methods with Simulation*. (Cambridge, UK: Cambridge University Press, 2003).

Yue, C. and Tong, C. 'Organic or local? Investigating consumer preference for fresh produce using a choice experiment with real economic incentives', *HortScience*, Vol. 44, (2009) pp. 366–371.