Food Quality and Preference 135 (2026) 105708

Contents lists available at ScienceDirect

Food Quality and Preference

journal homepage: www.elsevier.com/locate/foodqual

Willingness to pay for animal welfare across labels, products, consumers, and time

Malte Oehlmann [a], Irina Dolgopolova [b], Christina M. Neubig [a], Jutta Roosen [a,*]

[a] Technical University Munich, TUM School of Management, Chair of Marketing and Consumer Research
[b] Maastricht University, Department of Marketing and Supply Chain Management

ARTICLE INFO

Keywords:
Willingness to pay
Animal welfare
Multi-level label
Choice experiment

ABSTRACT

In recent years, the number of labels indicating improved animal welfare conditions on meat products has increased making it difficult for consumers to understand, evaluate, and compare husbandry conditions across products. Based on a discrete choice experiment implemented in three cross-section surveys over a period of 15 months with a total of 6000 German respondents, we estimate the willingness to pay (WTP) for various levels of animal welfare associated with different meat products. We use three existing labels with overlapping animal welfare requirements mimicking the situation in the German meat market: The well-established organic label as well as a binary animal welfare label by the Animal Welfare Initiative and a multi-level animal husbandry label which were introduced in Germany in 2015 and 2019, respectively. We show that the multi-level label scheme leads to more product differentiation and, subsequently, higher WTP estimates. WTP further depends on meat type, where animal welfare improvements for beef and chicken products are valued much higher compared to those for pork. WTP for the organic and the highest level of the husbandry label increases with higher household incomes. WTP for these labels on chicken is also higher among women.

1. Introduction

In recent years, meat companies and retailers have begun to introduce labels indicating improved levels of animal husbandry conditions. These labels are a response to the rising consumer demand for more transparency and better knowledge about the conditions under which animals were raised (Clark et al., 2017; Lagerkvist & Hess, 2011). This also holds for Germany, where – similarly to other countries – the number of food labels has increased. This is particularly true for labels indicating animal husbandry conditions and related animal welfare. Short of a unified introduction of a public labeling scheme by the German government, various animal welfare labels were introduced by private organizations.

As an initiative of partners from agricultural organizations, the meat industry, and the retail sector, the Animal Welfare Initiative (AWI) introduced a label in 2015. The label was applied to meat and meat products from participating farms implementing various measures to improve animal welfare, including extra space, toys to avoid boredom, and/or outdoor access. However, the label was meant to support the improvement of animal welfare conditions in an evolutionary process, so that improvements were conditioned on the farm-level situation when joining the initiative. As a result, the label does not allow for discerning the level of husbandry conditions that apply to the specific animal product (Initiative Tierwohl, 2025). In consequence, the label was criticized by consumer organizations for lacking transparency, as consumers cannot judge the actual welfare conditions under which the animal was raised.

Therefore, the German Society for the Promotion of Animal Welfare – a company developed by partners of the AWI – introduced a husbandry labeling scheme in 2019. The husbandry labeling scheme distinguishes four levels of animal husbandry for various meat products. The first level of the label stands for husbandry conditions that fulfill legal requirements, followed by level 2, which adds extra space, level 3, which allows access to outdoor climate, and finally level 4, where husbandry conditions require even more space and outdoor access, similar to the organic label.[1]

* Corresponding author.
E-mail address: jroosen@tum.de (J. Roosen).

[1] Since August 2023 a law passed by the German government introduces an animal husbandry label with 5 levels, where outdoor access and organic are differentiated into level 4 and 5. The label becomes mandatory for pork from animals raised and slaughtered in Germany in spring 2026. As a result of the new regulation, the private multilevel animal husbandry label has been adjusted to five levels as well.

https://doi.org/10.1016/j.foodqual.2025.105708
Received 12 March 2024; Received in revised form 3 September 2025; Accepted 13 September 2025
Available online 19 September 2025

These animal welfare labeling schemes were introduced in addition to the German organic label, which has existed since 2001. As illustrated in Appendix A, at the example of pork, the four-level husbandry labeling scheme, the participation label of the AWI, as well as the organic label are characterized by a lack of consistency and overlapping requirements regarding husbandry conditions. Consequently, it can be challenging for consumers to use these labels to understand, evaluate, and compare the level of animal welfare. Marette (2014) shows that label proliferation is an issue that can reduce consumer welfare.

In this study, we are interested in understanding how consumers react to information from several labels with redundant and/or conflicting information, as well as price, in their food choices. We compare willingness to pay (WTP) estimates for the differentiated multi-level husbandry scheme to the more general labels, such as the organic label, simultaneously accounting for different meat and meat replacement products, i.e., chicken, pork, beef, and vegetarian options. Since WTP for products labeled with animal welfare information might evolve over time and is subject to changes in consumer knowledge of the labels and prices, we measure WTP in three independent survey waves over a period of 15 months. Our study uses data from a large-scale discrete choice experiment (DCE), which explicitly allows us to account for trade-offs between labels as well as price in a multi-product context. Each of the three survey wave collected data from 2000 respondents each, yielding a total sample of 6000 consumers.

The remainder of this paper is structured as follows: The next section provides the theoretical background for this paper. Section 3 details the study design, the sampling, the choice experiment design, and the modeling approach. Descriptive statistics, as well as model results, are presented in Section 4 before this paper ends with a discussion of our findings and concluding remarks.

2. Theoretical background

Our research is predominantly based on two strands of literature: First, the literature on preferences for farm animal welfare, and second, the literature on trade-offs between food labels with redundant or overlapping information. Within the first stream of literature, Boaitey and Minegishi (2020) conducted a systematic literature review on the impact of consumers' socio-demographic characteristics, ethics and attitudes, product characteristics, and public roles that mattered for identifying animal welfare-conscious consumers. They found that individuals with higher levels of education and income, as well as women, consistently preferred higher animal welfare standards. Many of these effects were context- and species-specific. In a meta-analysis, Yang and Renwick (2018) investigated consumers' WTP a price premium for credence attributes of livestock products. Compared to products with 'mixed' attributes (e.g., attributes were described as 'natural' or 'good'), organic products were estimated to be associated with the highest price premium, followed by products bearing indications of hormone/antibiotic-free, animal welfare, food safety, and geographical indications. WTP further varied by product, with beef and dairy products being associated with a higher price premium compared to lamb. In their meta-analysis, Clark et al. (2017) found that consumers are willing to pay a small price premium for farm animal welfare, equivalent to just over half a standard deviation. The lowest average WTP estimate was observed for pigs, and the highest for beef and dairy products. Furthermore, WTP decreased with age, increased with income, and was higher for women, with both age and income having statistically significant effects. Similar effects of income and age were observed in a meta-regression by Lagerkvist and Hess (2011). In contrast to Clark et al. (2017), Lagerkvist and Hess (2011) observed that WTP for animal welfare was independent of the species from which the food product originated.

Considering the difficulties in measuring animal welfare and health outcomes (Dusel & Wieck, 2023), the vast majority of studies investigating the WTP for animal welfare focused on specific aspects of animal housing conditions and other production aspects that affect animal welfare such as outdoor access and stocking density (Latacz-Lohmann & Schreiner, 2018, de Jonge et al., 2015), castration and the transportation of animals (Grunert et al., 2018; Lagerkvist et al., 2006), or husbandry conditions during the rearing process of young animals (Gross et al., 2021). Typically, these studies analyzed preferences for specific animal types – see Frey and Pirscher (2018), who studied various aspects for chicks, hens, piglets, and pigs for one exception – with pigs being looked at most frequently (Clark et al., 2017).

The literature evaluating different labels considers trade-offs between products labeled as organic and/or local (e.g., Hasselbach & Roosen, 2015; Meas et al., 2015) or between an organic, a regional, and a climate label (e.g., Feucht & Zander, 2018). Particular attention has been paid to how consumers perceive and evaluate labels that provide redundant (Wilson & Lusk, 2020), overlapping, or contradictory information (Bernard et al., 2019). In this light, a study by Sonntag et al. (2023), who implemented a DCE in the German context, considered an animal welfare label, a climate label, an organic label, and the Nutri-Score for chicken breast and whole milk. They found that the presence of a label does not diminish the marginal utility of other labels. Respondents could handle two different types of labels simultaneously and seemed to be able to cope even with contradictory information in a trade-off situation between various labels. Schwickert (2023) evaluated a proposal for an animal welfare label in Germany (different from the one finally introduced) in combination with the organic label using a choice experiment and found that the three star (highest) level of the animal welfare label combined with the organic label would command a premium that was higher than the premium for each of the individual labels. Akaichi et al. (2022) also evaluated an animal welfare label in presence of an organic label. In a latent class analysis of a choice experiment with Scottish consumers, they discovered a class of consumers of 20 % whose WTP for the animal welfare label could be significantly increased when the organic label was also present.

Consumers' evaluation of multi-level labels, such as the husbandry label or the frequently employed Nutri-Score, which combines a hierarchical ranking of nutritional qualities with a letter and color scheme, have been the explicit focus of recent research. Weinrich and Spiller (2016a) compared a binary to a multi-level animal welfare label. Using an open-ended WTP question, they found that a multi-level labeling system achieved higher market shares than a binary label. Based on the same dataset, Weinrich and Spiller (2016b) observed significant differences for seven out of ten statements building the latent variable of satisfaction. They concluded that consumers might be more satisfied with a multi-level label. Schwickert (2023) evaluated an animal welfare label using one to three stars and the German organic label in a DCE using unprocessed (Schnitzel) and processed (sausage) pork and found a positive WTP for all levels of the animal welfare label. Carlsson et al. (2022) used a DCE with ready-cooked lasagna to evaluate label design using traffic light colors, grey color, or plain text to give information on antibiotics use, animal care, climate impact, and product healthiness. They found that the colored version of the label increases consumers' WTP to go from an inferior level (red) to an adequate level (yellow) and that the marginal utility to a higher level (green) decreases. Werle et al. (2022) showed that more levels help consumers to better assess and value the healthiness of products in the context of a Nutri-Score label.

Lagerkvist et al. (2014) considered the role of contextual dependence in choice experiments on attribute importance by studying trade-offs between origin and quality cues in a choice experiment with and without price. They conclude that the consideration of price leads to higher choice heterogeneity in the sample. However, the relative ranking of attribute importance and WTP remains similar.

Our research makes two major contributions to the above-mentioned literature: First, we analyze WTP for animal husbandry conditions over time using three survey waves, controlling for possible influences by respondents' label knowledge and trust as well as their socio-demographic characteristics. Studies on WTP over time are still very

scarce in the food choice literature and are mostly discussed in the context of test/retest reliability (e.g., Rigby et al., 2016, Mørkbak & Olsen, 2015).

Our study was inspired by Lusk and Tonsor (2016), who conducted a DCE among U.S. consumers over one year with twelve survey waves in total. Their focus was not on food labeling but on elasticities and substitution patterns between different meat cuts. As they do, and in contrast to Bolos et al. (2022), we draw fresh samples in each of our surveys and do not follow a repeated measure design. This research design allows us to study the stability of preferences for animal welfare characteristics over time in a phase of high inflation.

Second, we add to the growing research on trade-offs between food labels. Given the influence of context, the DCE is inspired to mimic product choice in a retail store, using labels as they are actually implemented on meat packages. In particular, we test how WTP reacts to the scope and levels of the husbandry labeling scheme and compare these WTP for multi-level labels to WTP for binary labeling schemes (AWI and organic).

3. Material and methods

3.1. Study design

This research is based on three online surveys with separate samples. A general description of the survey waves can be found in Christoph-Schulz et al. (2024). Besides the collection of socio-demographic information, each survey included a variety of questions, such as statements on eating behavior, attitudes towards animal welfare issues in agriculture, etc. Prior to the DCE, participants were asked to state whether they knew and trusted a variety of food labels, including those considered in the DCE. While respondents' knowledge was asked on a binary scale (yes or no), trust in the labels was indicated on a 5-point Likert scale. Trust was not measured for labels that were unknown to the respondent.

3.2. Survey sampling

Data was collected in three separate surveys, each time using a fresh sample of respondents. The design of the DCE stayed the same. Survey 1 took place from February to April 2022, Survey 2 was carried out from July to September 2022, and Survey 3 took place in April 2023. Market research companies were employed for programming, sampling, and implementing the surveys. To ensure representativeness, quotas with respect to gender, age, region (north, south, east, and west), and monthly net household income were applied to make the samples match the general German population aged 18–70.

Respondents who always chose the same position across all nine choice sets, e.g., always the third alternative, were excluded from the analysis. This was the case for one respondent in Surveys 1 and 3, as well as for five respondents in Survey 2. As the order of the products was randomized, it is assumed that these participants did not give valid answers. However, people who always chose the "no choice" option were considered for the analysis. The final sample thus includes about 2000 subjects per survey, making for a total of 5993 respondents.

Table 1 summarizes the descriptive statistics regarding gender, age, and monthly net household income. Overall, the percentages are relatively stable across the three survey waves, with only a slight increase in female respondents (decrease in male participants) throughout the study period. Under the null hypothesis that the observed frequencies match the expected frequencies, a χ^2-test ($\chi^2(4) = 2.95$, $p = 0.57$) provides evidence that the number of respondents by sex is independent of the survey wave. The age distribution is rather stable in Surveys 1 and 2, but differs with regard to the oldest age group in Survey 3. Overall, age distribution depends on survey wave ($\chi^2(10) = 23.58$, $p = .009$). Regarding the income distribution, it appears that the two lowest income groups are less frequent in Survey 3 compared to Surveys 1 and 2, whereas the opposite is observed for the middle-income group. Especially in Survey 3, the proportion of the highest income group is substantially higher, and the proportion of the second lowest income group is lower compared to the first two surveys. A χ^2-test ($\chi^2(8) = 90.02$, $p < .001$) indicates that the household income depends on the survey wave.

Table 1
Socio-demographic characteristics of participants across surveys.

Characteristics	Frequencies in %			
	Survey 1 (N = 1999)	Survey 2 (N = 1995)	Survey 3 (N = 1999)	Germany
Gender				
Female	50.0	50.5	51.2	50.7
Male	49.9	49.4	48.6	49.3
Diverse	0.1	0.2	0.3	n/a
Age group (years)				
18–23	4.1	3.9	3.6	9.1
24–33	16.8	16.3	19.3	18.2
34–43	18.7	18.3	18.8	18.3
44–53	18.1	18.2	18.8	19.5
54–63	23.2	23.9	24.8	22.8
64–70	19.1	19.4	14.7	12.1
Monthly net household income				
< 1500 Euros	26.0	24.9	21.5	25.7
1500–2599 Euros	31.5	32.1	23.9	31.1
2600–3199 Euros	11.9	12.7	16.4	11.7
3200–4499 Euros	16.7	17.4	18.7	16.6
≥ 4500 Euros	14.0	12.8	19.5	14.9

3.3. DCE design

The statistical design of the DCE was inspired by Lusk and Tonsor (2016) and was constant across the three surveys. Each respondent saw nine choice sets, and each choice set was composed of nine alternatives, including the no-choice option. These nine alternatives were intended to realistically represent the variety of products available to consumers in Germany when shopping for a meal. As a consequence, both meat and vegetarian alternatives were considered. Petersen et al. (2024) estimate, based on retail scanner data, that meat substitutes make up about 9 % of sales in relevant meat categories. Typical products were selected for the different animal species based on information from a large scanner dataset provided by the Society for Consumer Research (GfK) in Germany. We consider the following alternatives:

- Beef minute steak (abbreviated as beef in the following)
- Chicken breast (abbreviated as chicken in the following)
- Minced pork (abbreviated as pork in the following)
- Vegetarian schnitzel (abbreviated as vschnitzel in the following)
- Vegetarian mincemeat (abbreviated as vmince in the following)

Each choice set included two different chicken, pork, beef, and vegetarian alternatives.

Based on the literature review and the market situation in Germany as detailed in the introduction, three labels were identified for inclusion in the choice experiment:

3.3.1. German organic label (hereafter referred to as organic)

The German organic label is a state label that companies can use voluntarily for organic food and agricultural products. It has existed since 2001 and can be applied in addition to the EU's organic label, which has been mandatory for prepackaged organic foods from the EU since July 2012. The label represents certified organic production and species-appropriate treatment of animals, e.g., no chemical growth regulators or hormones, low stocking density, organic feed, etc. (BLE

Projektgruppe Ökolandbau, 2023; BMEL, 2022).

3.3.2. Label of the animal welfare initiative (AWI)

The animal welfare initiative (in German: Initiative Tierwohl) is an association of partners from agriculture, the meat industry, food retailers, and gastronomy that has existed since January 2015. They developed a label for meat products (poultry and pork), indicating that the meat originates from a farm or company that is part of the initiative. Participating farms must fulfill specific criteria, including hygiene standards, monitoring of antibiotics, a healthy climate, daylight, more space, and veterinary care (Initiative Tierwohl, 2025). Criteria and the corresponding monetary compensation vary across farmers.

3.3.3. Husbandry label (hereafter referred to as husbandry)

The AWI was also involved in the development of a husbandry labeling scheme for packaged meat, which has existed for beef, chicken, and pork (among others) since April 2019, introduced by the German Society for the Promotion of Animal Welfare in Farm Animal Husbandry (Gesellschaft zur Förderung des Tierwohls in der Nutztierhaltung, 2023). Products are categorized into four levels – each visualized by a color – which differ in terms of stocking density and whether the animals have access to the outside climate (Initiative Tierwohl, 2019):

- Husbandry 1 "stable/indoor housing": Legal minimum standards are met
- Husbandry 2 "stable housing plus": Lower stocking density compared to Husbandry 1, which corresponds to the standards of the AWI label
- Husbandry 3 "outdoor access": Animals have access to the outside space, lower stocking density as for Husbandry 2
- Husbandry 4 "premium": Lower stocking density compared to Husbandry 3, access to outside space, fulfills husbandry requirements of the organic label

Each alternative was characterized by the presence or absence of a pictogram of these three labels. The overview of labels is presented in Table 2.

Since the label for husbandry level 1 is hardly present in the market, only labels for husbandry levels 2, 3, and 4 were used. To derive product prices, different grocery store chains were visited and online sources were evaluated prior to each survey wave. Despite high food price inflation in 2022 and 2023, no major price changes were observed during the study period for the products in question. Hence, the price vector, which contained current market prices, was held constant across all surveys. The price could adopt six different levels, which were defined on a product-specific basis and each related to a product of 400 g. Following regulatory requirements (Directive EC98/6), the price per kilogram was shown as well. Table 3 depicts the attributes and their levels for the respective products.

To assign attribute levels to the choice sets, a d-efficient design with flat (zero) priors was implemented using the software Ngene (Ngene, 2018). In this way, three blocks, each having nine choice sets, were created. During the survey, each respondent was randomly assigned to one of the three blocks. Within each block, the order of the choice sets was randomized. Furthermore, the order in which the product alternatives were shown was randomized for each choice set. An exception was the "no choice" alternative, which was always presented last. The design included several constraints. For example, an organic product was either shown with "Husbandry 4" or no husbandry label, but not with "Husbandry 2" or "Husbandry 3".

Each choice set was introduced with the following sentence: "Imagine you are at the grocery store buying the ingredients to prepare a meal for you or your household. For each of the nine questions that follow, please indicate which product you would be most likely to buy." Fig. 1 gives an example of a choice set.

Table 2
Pictograms of the labels used in the choice experiment.

Label name	Pictogram
German organic label	
Label of the animal welfare initiative	
Husbandry 1 "stable housing"	
Husbandry 2 "stable housing plus"	
Husbandry 3 "outdoor access"	
Husbandry 4 "premium"	

Table 3
Attributes and levels.

Product	Organic	AWI	Husbandry	Price (400 g) in Euros
beef	no; yes		0; 2; 3; 4	3.58; 3.98; 4.38; 6.27; 11.14; 16.00
chicken	no; yes	no; yes	0; 2; 3; 4	2.18; 2.58; 2.98; 5.08; 8.52; 11.56
pork	no; yes	no; yes	0; 2; 3; 4	1.93; 2.20; 3.19; 6.53; 7.25; 7.96
vschnitzel	no; yes			6.27; 6.96; 7.44; 9.77; 10.57; 11.37
vmince	no; yes			2.79; 4.06; 5.32; 6.52; 7.08; 7.36; 7.36

Note: 0 = Label was not presented.

3.4. Modeling approach

The data obtained from the DCE is analyzed using the random utility model (McFadden, 1973). For respondent n and alternative j=1, 2, …, 9 in choice situation t, we specify the utility function in WTP-space as follows:

$$U_{n,j,t} = \sum_{s=1}^{4} \delta_{s,n,j} + \beta_{price,n} \left(price_{n,j,t} + \sum_{s=1}^{4} \left(\beta_{org,s,n,j} \times org_{s,n,j,t} + \beta_{AWI,s,n,j} \times AWI_{s,n,j,t} + \beta_{h2,s,n,j} \times h2_{s,n,j,t} + \beta_{h3,s,n,j} \times h3_{s,n,j,t} + \beta_{h4,s,n,j} \times h4_{s,n,j,t} \right) \right) + \epsilon_{n,j,t} \quad (1)$$

Here, $\delta_{s,n,j}$ are the product-specific constants where $s = 1$ stands for beef, $s = 2$ stands for chicken, $s = 3$ stands for pork, and $s = 4$ stands for the vegetarian options (vschnitzel and vmince). The "no choice"

Fig. 1. Choice set example (translated from the German original).

alternative was always indexed as $j = 9$. $price_{n,j,t}$ is the price of alternative j (0 for the "no choice" option), $org_{s,n,j,t}$ is a dummy variable that takes a value of 1 if the organic label was shown on alternative j for product type s, 0 otherwise. Therefore, we estimate label effects per product type (beef, chicken, pork, vegetarian), as the literature suggests that WTP for animal welfare is species dependent. Similarly, $AWI_{s,n,j,t}$ is a dummy variable that adopts a value of 1 if the label of the animal welfare initiative was shown on alternative j and product type s, 0 otherwise. $AWI_{s,n,j,t}$ was always 0 for beef and the vegetarian options as well as the "no choice" option, as the label exists only for chicken and pork. $h2_{s,n,j,t}$, $h3_{s,n,j,t}$, and $h4_{s,n,j,t}$ are the dummy variables for the second, third, and fourth level of the husbandry labeling scheme, which adopt a value of 1 if the corresponding label was shown, 0 otherwise. These dummies were always 0 for the "no choice" alternative. $\epsilon_{n,j,t}$ is a type I extreme value error term that is distributed identically and independently across alternatives and observations.

To arrive at a final model specification, the first step was to estimate simple multinomial logit (MNL) models for each survey separately as well as the pooled sample of the three surveys. Then we proceed to identify random choosers as suggested by Malone and Lusk (2018). To this end, we estimate a latent class model with all parameters restricted to zero in one class using lclogit2 (Yoo, 2020) in Stata. The parameter vector in the second class was unconstrained. Identifying class membership based on class probability, the subsequent analyses are done using only nonrandom choosers. These are 1218 (60.9 %) in Survey 1, 1268 (63.6 %) in Survey 2, and 1214 (60.7 %) in Survey 3. The following results are based on these 3700 respondents.

After convergence of all models in preference space, we relaxed the assumptions of the MNL and estimated Mixed Logit (MXL) models (Train, 2009) for each survey wave in preference as well as WTP space using the mixlogit package in Stata (Hole, 2007). Since we are interested in WTP estimates, the final model is a MXL in WTP space with 250 Halton draws. To allow for unobserved preference heterogeneity, the price coefficient is assumed to follow a log-normal distribution. All other parameters, including the product-specific constants, are assumed to be normally distributed.

We obtain posterior estimates of WTP for each respondent and merge the data from the three survey waves. We use regression analysis to investigate whether WTP changes across the three survey waves while controlling for the impact of respondents' label knowledge and trust, as well as their socio-demographic characteristics. The posterior means were used as the dependent variable in an ordinary least-squares (OLS) regression, which was specified as follows:

$$WTP_{n,k} = \gamma_k + b_{w2,k} \times w2_n + b_{w3,k} \times w3_n + b_{lt,k} \times tl_{n,k} + b_{ht,k} \times th_{n,k} + b_{fem,k} \times fem_n + b_{age,k} \times age_n + b_{inc15,k} \times inc15_n + b_{inc26,k} \times inc26_n + b_{inc32,k} \times inc32_n + b_{inc45,k} \times inc45_n + e_{n,k} \quad (2)$$

Here, n refers to the respondent in the dataset from all three surveys. Therefore, for each of these regression models, the number of observations was equal to $N = 3700$ respondents.

The index k corresponds to one of the 14 product-label-specific WTP estimates (e.g., Chicken organic, Chicken Husb. 2, etc.). No OLS regressions were conducted for the product-specific constants, the price coefficient, or the organic label for the vegetarian products. $w2$ and $w3$ are dummy variables indicating the survey waves 2 and 3, respectively. As we continue to tl, it is a dummy variable that takes a value of 1 if respondent n knows the label k and stated a low level of trust, i.e., ≤ 3 on the 5-point rating scale. th takes a value of 1 if respondent n knows the label k and stated a high level of trust, i.e., a value of 4 or 5 on the 5-point rating scale. fem takes a value of 1 if respondent n is female, 0 otherwise. age is the age of respondent n in years. $inc15$ takes a value of 1 if respondent n indicated to have a net household income between 1500 and 2599 Euros, 0 otherwise. $inc26$ takes a value of 1 if respondent n indicated to have a net household income between 2600 and 3199 Euros, 0 otherwise. $inc32$ takes a value of 1 if respondent n indicated to have a net household income between 3200 and 4499 Euros, 0 otherwise. $inc45$ takes a value of 1 if respondent n indicated to have a net

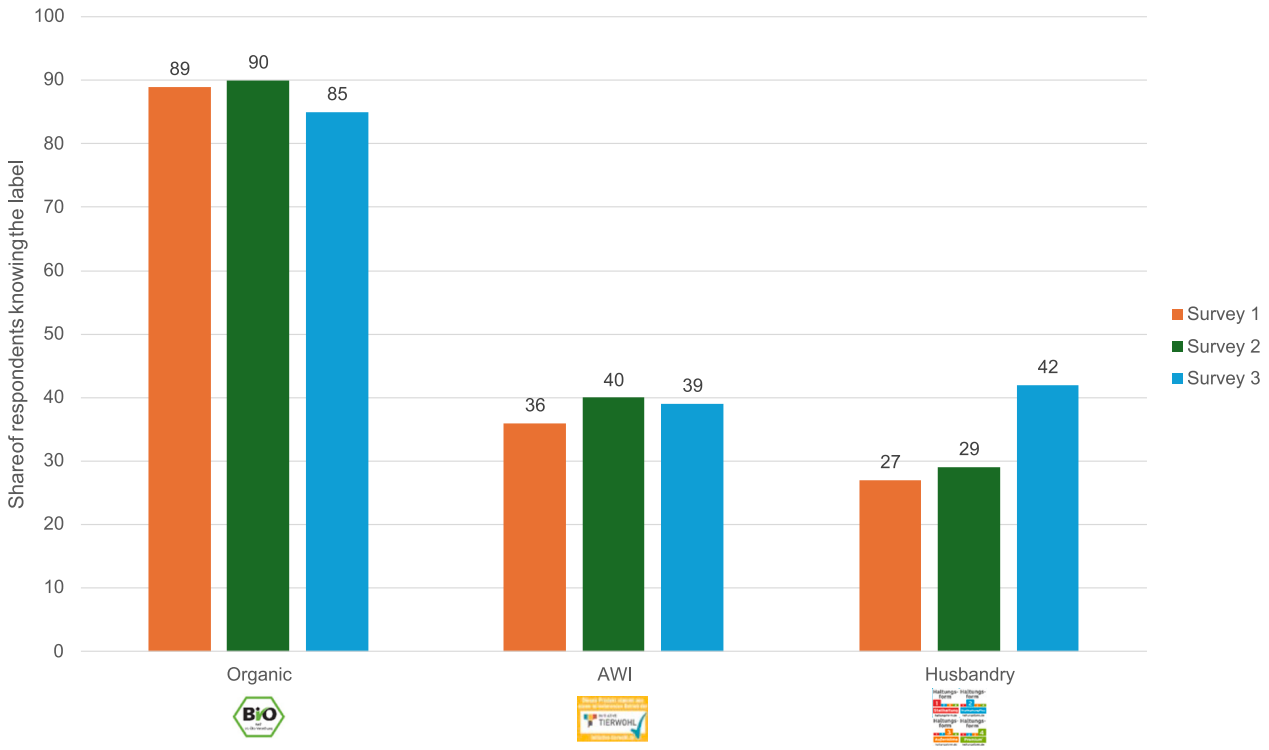

Fig. 2. Respondents' label knowledge.

Fig. 3. Mean trust in labels known to the respondents.

household income between larger than 4500 Euros, 0 otherwise.

4. Results

4.1. Descriptive statistics

Fig. 2 summarizes respondents' answers on label knowledge – collected prior to the DCE on a binary scale (yes or no) – for the three labels used in the DCE. Participants state, on average, the highest knowledge of the German organic label. Eighty-five percent or more of the respondents know the label. The percentage of respondents who either know the AWI label or the labels of the husbandry scheme is much lower. While the knowledge of the organic label is rather stable during the study period, a significantly higher knowledge of the AWI label can be observed for Survey 2 and 3. Interestingly, the knowledge of the husbandry label scheme increased considerably in Survey 3 compared to the earlier surveys. As this is the label most recently introduced, we would expect label knowledge to increase over time.

Respondents were subsequently asked to state their trust in the label on a five-point Likert scale if they indicated knowing the label (Fig. 3). Across labels, average trust in the organic label was highest, although the differences in mean trust are only small. Across survey waves, variations in the means are very minor, with Survey 1 having slightly higher numbers. The variables knowledge and trust were transformed into the dummy variables *low trust* (know label and trust ≤ 3) as well as *high trust* (know label and trust > 3).

Turning to the descriptive analysis of the data obtained from the DCE, Table 4 presents the relative frequency of the chosen alternatives by product and survey (only non-random choosers). Across all three surveys, chicken is the product selected most frequently, accounting for about half of the choices. The alternatives beef and pork were each chosen with similar frequency of 22–25 %. With around 1 %, the proportion of instances in which the vegetarian products were chosen is much lower. The frequency of choices in which non-random choosers chose none of the eight products offered on the choice set ranges from 3.82 % (Survey 1) to 2.55 % (Survey 3). The proportion of the "no choice" alternative decreases across all three surveys.

4.2. Willingness to pay across labels and products

Table 5 depicts the results of a MXL estimated in WTP space per survey wave. Besides the means and standard deviations for the marginal WTP estimates per label, which are expressed in Euros for a product of 400 g, the corresponding estimates for the alternative-specific constants (beef, chicken, pork, and vegetarian), as well as the price coefficient, are shown. The vast majority of the effects have the expected sign and are statistically significant.

The estimates for the product-specific constants, which need to be compared to the "no choice" alternative, are sizeable. This means that the participants in the estimation sample value the proposed products. The magnitude and standard errors of their estimated standard deviations indicate a high degree of heterogeneity with respect to the product-specific effects, which are not captured by the labels. The mean of the price parameter has the expected negative impact on respondents' utility (around −1.7) with a rather small standard deviation (between 0.08 and 0.17). Both are highly significant.

Table 4
Relative frequency in percentage of chosen alternatives across surveys.

Alternative	Survey 1 (10,962 choices)	Survey 2 (11,412 choices)	Survey 3 (10,926 choices)
Beef	23.22	21.82	22.49
chicken	48.97	48.30	48.53
pork	22.44	25.49	24.59
vschnitzel	0.52	0.46	0.75
vmince	1.02	0.92	1.09
no choice	3.82	3.01	2.55
Sum	100	100	100

Turning to the WTP estimates and looking first at the organic label, the highest WTP is found for the vegetarian products (9.98 / 8.22 / 12.44 Euros), followed by the average WTP for the beef minute steak (5.05 / 4.56 / 4.82 Euros), chicken breast (2.82 /2.15 /1.98 Euros), and minced pork (not significantly different from 0). Large and significant standard deviations reveal a high degree of heterogeneity in the sample. For the AWI label, which is only used on chicken and pork, WTP estimates are of the same magnitude for chicken and pork. The values are 2.82 versus 2.62 in Survey 1, 2.02 and 1.51 in Survey 2, and 1.79 and 1.78 in Survey 3 for chicken and pork, respectively. Compared to the organic label, not only the means but also the standard deviations are much smaller and often not significant, suggesting rather homogeneous preferences.

WTP continuously increases over the four levels of the husbandry labeling scheme. This increase is almost linear as the level increases from 2 to 3 and from 3 to 4. Across all products and labels, the highest WTP is observed for the fourth husbandry level. Here, respondents are willing to pay 10.74 /10.45 /10.26 Euros (beef in surveys 1, 2, and 3, respectively), 7.82 /7.04 /5.91 Euros (chicken), and 8.31 /7.25/ 7.50 Euros (pork) if the husbandry label level 4 is present on the product. Although animal welfare standards for this label are equivalent to those of the organic label, WTP is much higher when animal welfare is made explicit through the husbandry label. However, the magnitude of the standard deviations suggests that this finding does not apply to the sample as a whole.

Fig. 4 shows kernel graphs of the WTP distributions for the organic label and the husbandry label level 4 for beef, chicken, and pork. The corresponding graphs for the product-specific constants, the AWI label, and the other levels of the husbandry label can be found in Appendix B. While the WTP distributions are stable across survey waves for the organic label, they are less so for the other labels that have been on the market for a shorter period.

4.3. Willingness to pay across consumers and time

Table 6 shows the results of the OLS estimation in which the dependent variable is the mean of the individual-specific WTP per label. Table 6 illustrates the results by taking the example of chicken breast. The corresponding estimates for beef and pork are qualitatively similar and can be found in Appendix C.

From Table 6, a clear pattern emerges: WTP for all labels except the husbandry label level 2 decreases as time progresses (from Survey 1 to Survey 2 and to Survey 3). Trust has a negative effect on WTP for all labels but not for the husbandry label level 2. The effect is stronger for high trust compared to low trust. It seems that while trust in labels may be low, labels still serve as an indicator guiding consumers' choices. Females have a higher WTP for the organic label and the husbandry level 4 label, but a lower WTP for the husbandry level 2 label. It seems that women value organic and high animal welfare standards and discount low standards. There is a clear income gradient for WTP for organic and husbandry level 4. Age has a significant positive effect on WTP for the husbandry level 3 label, but has no significant effect for any other label.

Regarding beef and pork, we observe similar income gradients for beef and pork and similar gender effects for pork (see Appendix C). Age is insignificant in most regressions. High trust has a positive effect on WTP for pork for the organic label, the AWI label, and the husbandry label level 4.

5. Discussion

Our WTP model results reveal that the use of multi-level animal welfare labels, where each level has a separate set of standards, can lead to more differentiated markets and, subsequently, higher WTP. The highest price premiums are observed for the fourth level of the

Table 5
Mean and Standard Deviation (Std. Dev.) of WTP in Euros (estimates by survey – MXL estimates in WTP space).

	Survey 1		Survey 2		Survey 3	
Variables	Mean	Std. Dev.	Mean	Std. Dev.	Mean	Std. Dev.
Beef	1.614***	6.546***	2.669***	6.447***	4.323***	7.445***
	(0.553)	(0.459)	(0.548)	(0.425)	(0.613)	(0.535)
Chicken	9.519***	5.394***	11.13***	5.008***	13.56***	4.561***
	(0.483)	(0.373)	(0.516)	(0.372)	(0.635)	(0.445)
Pork	3.061***	5.855***	5.535***	6.092***	6.734***	6.458***
	(0.503)	(0.457)	(0.486)	(0.397)	(0.560)	(0.444)
Vegetarian	−10.37***	5.391***	−6.993***	3.270***	−9.263***	5.001***
	(1.377)	(0.865)	(1.230)	(1.223)	(1.481)	(0.970)
Organic (beef)	5.051***	12.18***	4.559***	12.16***	4.817***	11.60***
	(0.599)	(0.675)	(0.603)	(0.621)	(0.626)	(0.639)
Organic (chicken)	2.816***	8.695***	2.146***	7.286***	1.975***	7.208***
	(0.393)	(0.497)	(0.365)	(0.455)	(0.390)	(0.530)
Organic (pork)	0.867	9.987***	0.632	9.142***	0.654	9.276***
	(0.591)	(0.830)	(0.555)	(0.741)	(0.617)	(0.870)
Organic (vegetarian)	9.984***	1.301	8.215***	−1.430	12.44***	1.039
	(1.165)	(1.144)	(1.106)	(1.199)	(1.306)	(1.430)
AWI (chicken)	2.801***	2.275***	2.021***	1.205	1.787***	0.895
	(0.341)	(0.708)	(0.307)	(0.845)	(0.317)	(0.761)
AWI (pork)	2.618***	0.758	1.511***	1.157*	1.775***	0.270
	(0.431)	(0.690)	(0.381)	(0.656)	(0.399)	(0.504)
Husbandry 2 (beef)	0.795	4.207***	2.476***	2.672**	1.072	5.081***
	(0.880)	(1.190)	(0.705)	(1.201)	(0.959)	(1.264)
Husbandry 2 (chicken)	−0.0646	3.165***	0.521	4.496***	−0.0744	3.320***
	(0.348)	(0.631)	(0.344)	(0.493)	(0.352)	(0.605)
Husbandry 2 (pork)	−0.129	8.126***	1.035*	7.063***	2.311***	6.333***
	(0.723)	(0.881)	(0.565)	(0.685)	(0.576)	(0.730)
Husbandry 3 (beef)	6.289***	0.610	5.617***	1.874**	6.316***	3.841***
	(0.570)	(1.270)	(0.573)	(0.870)	(0.670)	(0.881)
Husbandry 3 (chicken)	3.764***	1.745**	3.463***	1.078	2.658***	1.940**
	(0.384)	(0.776)	(0.358)	(1.007)	(0.384)	(0.794)
Husbandry 3 (pork)	4.208***	0.892	3.967***	0.0159	3.924***	1.040
	(0.521)	(0.985)	(0.464)	(1.176)	(0.508)	(0.969)
Husbandry 4 (beef)	10.74***	6.356***	10.45***	4.587***	10.26***	5.438***
	(0.584)	(0.599)	(0.545)	(0.709)	(0.574)	(0.642)
Husbandry 4 (chicken)	7.820***	7.564***	7.042***	8.471***	5.908***	9.700***
	(0.460)	(0.559)	(0.442)	(0.549)	(0.473)	(0.604)
Husbandry 4 (pork)	8.309***	7.287***	7.250***	9.799***	7.497***	8.488***
	(0.586)	(0.784)	(0.583)	(0.753)	(0.577)	(0.799)
Price	−1.656***	0.0778**	−1.641***	0.127***	−1.708***	0.170***
	(0.0385)	(0.0327)	(0.0379)	(0.0304)	(0.0410)	(0.0315)
Number of choices	10,962		11,412		10,926	
Number of individuals	1218		1268		1214	

Standard errors in parentheses.
*** $p < 0.01$, ** $p < 0.05$, * $p < 0.1$.

husbandry labeling scheme, although this level is equivalent to the organic label in terms of animal husbandry conditions. This contrasts with the results by Schwickert (2023), who found similar WTP for the organic label and the highest level of the animal welfare label. A reason for this might be the coloring of the husbandry labels (red, blue, orange, green) as well as the numbering (1, 2, 3, 4) from which consumers can directly infer the order. Similar results were also obtained by Sonntag et al. (2023), Feucht and Zander (2018), and Thøgersen and Nielsen (2016), who found that participants preferred multi-level, multi-color labels.

In addition, multi-level labels might shift the reference point when evaluating meat products. Although level 2 of the husbandry scheme requires higher animal welfare standards compared to products not carrying this label, we find that WTP for this label is small or even not statistically significantly different from zero. In comparison, the AWI label, which guarantees the animal welfare standards of level 2 or higher of the husbandry scheme, and is thus more ambiguous, yields a price premium slightly larger but still comparable in magnitude to that of husbandry level 2. These results might, however, also be triggered by the design of our DCE, in which respondents likely saw husbandry label levels 2, 3, and 4 on the same choice set, which might have caused respondents to infer that the second level of the scheme entailed the lowest standard. Contrary to Carlsson et al. (2022), we don't observe declining marginal WTPs when raising animal husbandry label from level 3 to 4 compared to raising it from 2 to 3. According to our results, WTP increases almost linearly when going to higher animal welfare standards.

When comparing WTP across products, our results are mixed. We find that the labels differentiate the WTP for the products most for beef and least for pork. This hierarchy corresponds to absolute prices found in supermarkets. Interestingly, differences in WTP across products (beef, chicken, pork) are much higher for the organic label where the WTP for the organic label on pork is not statistically different from zero. However, these findings might have been at least partially affected by the design of our study in which not only the animal species but also the processing of the meat differed across alternatives, i.e., chicken breast versus minced pork versus beef minute steak. Staudigel and Trubnikov (2022) find in a hedonic price analysis based on scanner data that minced meat commands the lowest price premium for organic among all meat cuts. We find that product-specific differences in WTP are smaller for the fourth level of the husbandry scheme.

Overall, we find some differences in WTP over time, although with mostly declining WTP for the high animal welfare standard. Hempel and Roosen (2024) saw a rising importance of the price motive in food values in a similar period in Bavaria. This highlights the importance of considering the context in which food choices are analyzed. WTP for the organic label, that is known by almost all consumers, is much more

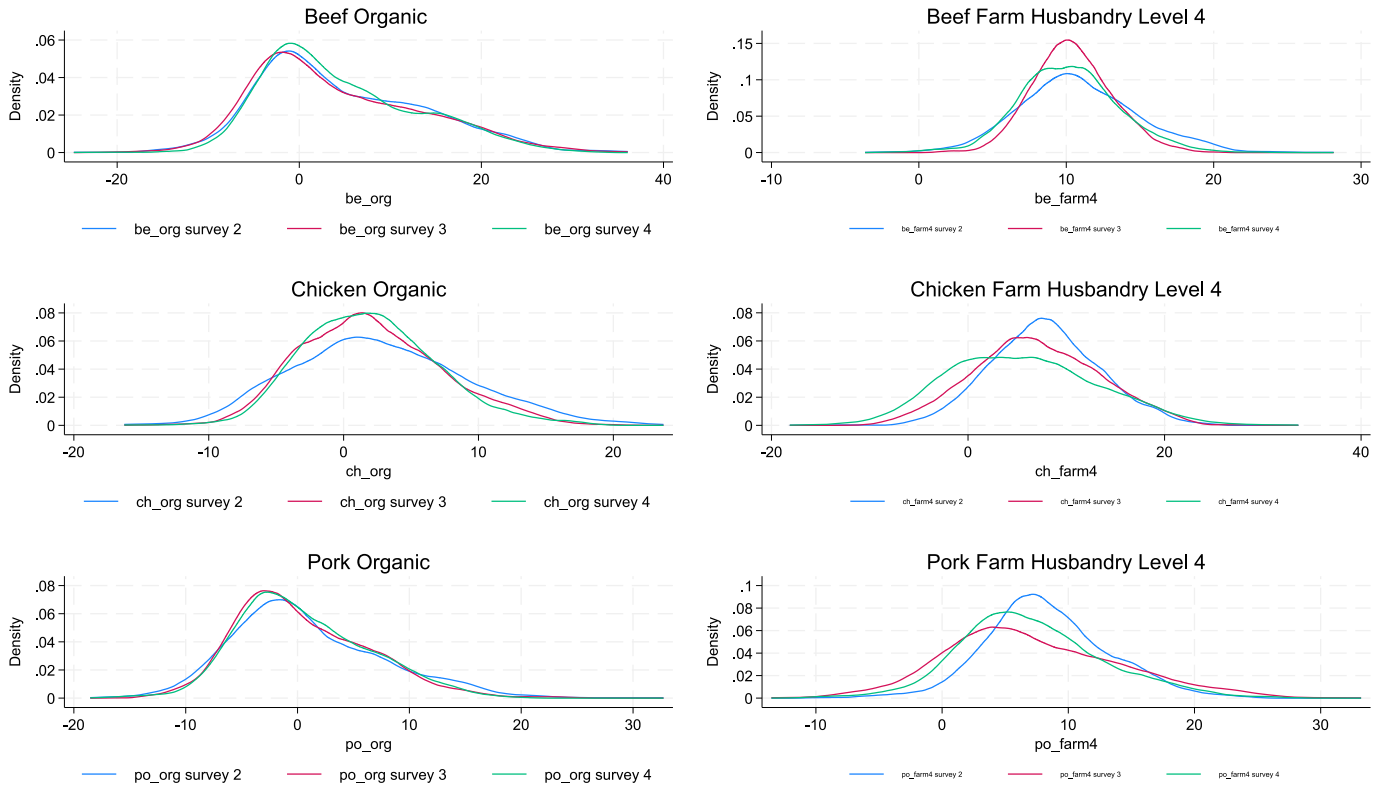

Fig. 4. Kernel density distribution for individual coefficients for the organic label and the husbandry level 4 label.

Table 6
WTP per label regressed on respondents' characteristics - Chicken.

Variable	WTP organic	WTP AWI	WTP Husb. 2	WTP Husb. 3	WTP Husb.4
Survey 1	Base				
Survey 2	−0.577***	−0.779***	0.641***	−0.334***	−0.827***
	(0.219)	(0.0290)	(0.0836)	(0.0293)	(0.256)
Survey3	−0.677***	−0.982***	0.0667	−1.127***	−2.455***
	(0.222)	(0.0294)	(0.0854)	(0.0300)	(0.261)
Don't know the label	Base				
Know and low trust	−0.577***	−0.779***	0.641***	−0.334***	−0.827***
	(0.219)	(0.0290)	(0.0836)	(0.0293)	(0.256)
Know and high trust	−0.677***	−0.982***	0.0667	−1.127***	−2.455***
	(0.222)	(0.0294)	(0.0854)	(0.0300)	(0.261)
Not female	Base				
Female	0.839***	0.0361	−0.175**	−0.0241	0.880***
	(0.180)	(0.0238)	(0.0688)	(0.0241)	(0.210)
Age	−0.0101	−0.000469	6.64e-05	0.00155*	−0.00502
	(0.00646)	(0.000850)	(0.00245)	(0.000861)	(0.00751)
Income <1500	Base				
Income 1500–2599	0.677***	0.00328	0.0613	−0.0312	1.072***
	(0.252)	(0.0334)	(0.0963)	(0.0338)	(0.295)
Income 2600–3199	0.644**	−0.0642	0.119	−0.0174	1.905***
	(0.310)	(0.0411)	(0.119)	(0.0416)	(0.363)
Income 3200–4499	1.308***	−0.0383	−0.132	−0.0341	2.102***
	(0.279)	(0.0369)	(0.106)	(0.0374)	(0.326)
Income ≥4500	1.738***	−0.0589	−0.259**	−0.00135	3.356***
	(0.299)	(0.0395)	(0.114)	(0.0400)	(0.349)
Constant	2.323***	2.813***	0.0826	3.728***	6.018***
	(0.422)	(0.0553)	(0.159)	(0.0560)	(0.488)
R-squared	0.038	0.257	0.026	0.292	0.054

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Income = Monthly net household income in Euros, N = 3700 observations.

stable compared to WTP for the animal welfare label, which was introduced only three years prior to the first survey of this study. Only the husbandry label level 4 yields results that are similar with regard to their stability in time.

With respect to the organic as well as the highest level of the husbandry label, WTP is positively affected by household income and gender (female). This finding is well in line with previous research (e.g., Boaitey and Minegeshi, 2020; Clark et al., 2017; Lagerkvist & Hess, 2011). In contrast to Clark et al. (2017) and Lagerkvist and Hess (2011), we do not find a consistent age effect. Also, respondents' awareness as

well as their trust seem to play a crucial role in driving WTP for animal welfare labels.

Our findings are of interest to marketers, food product retailers, and policymakers. If one were interested in increasing WTP for meat from farms with higher animal welfare standards, a strategy could be to make consumers aware and enhance their trust in the corresponding labels. However, as we see in the case of chicken, this does not necessarily increase WTP. Our study also indicates that breaking down complex concepts like animal welfare can simplify consumers' decision-making processes, especially when colorful, traffic light-like labels are used. This confirms results by Carlsson et al. (2022) and Werle et al. (2022). The simplification occurs when the consumers can infer the importance of different standards from the label. In our case, the generic label of the AWI commands a WTP comparable to that for the level 2 of the husbandry label. This WTP falls short of the price premia that high levels of the husbandry label can command. Our results also indicate that different products can be priced differently when it comes to the organic label.

5.1. Limitations

This study is not without limitations. First, it was not possible to fulfill all quotas to the same extent across survey waves. Second, it must be acknowledged that while the DCE stayed constant over time, the identification of nonrandom choosers led to a varying distribution of blocks in the samples. This may have led to some product label combinations being present less frequently compared to others, which may have affected the uncertainty around the mean WTP estimates.

Further research could have an explicit look at substitution patterns between meat and vegetarian alternatives that are conquering the markets. However, for the nonrandom choosers analyzed in this paper, the market share for vegetarian alternatives is small. A question that might arise is to what extent demand moves from meat-based to vegetarian products when products have an animal welfare label. Moreover, revealed-preference data could be used to test whether our findings, which were obtained through hypothetical choice scenarios, hold in the real world.

6. Conclusion

This study tracked the evolution of WTP for products bearing animal welfare labels over a period of 15 months. In three survey waves, we analyzed trade-offs between labels covering various aspects of animal welfare and the price of the product. The labels employed vary with regard to how explicitly they state animal husbandry requirements. Descriptive results indicate that respondents' knowledge of the organic label is much higher compared to the two labels explicitly aiming at animal welfare aspects. This is most likely due to the fact that the organic label was introduced much earlier (2001) compared to the label of the AWI (2014) and the husbandry labeling scheme (2019). Our results show that the multi-level label scheme leads to more product differentiation and, subsequently, higher WTP estimates. WTP further depends on meat type, where welfare improvements for beef and chicken products are valued much higher compared to welfare improvements for pork. WTP varies across the three survey waves. For the organic label and the highest level of the husbandry label, WTP increases with higher household incomes. Here, WTP is also higher among women.

CRediT authorship contribution statement

Malte Oehlmann: Writing – original draft, Visualization, Project administration, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Irina Dolgopolova:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Christina M. Neubig:** Writing – review & editing, Project administration, Investigation, Data curation. **Jutta Roosen:** Writing – review & editing, Supervision, Project administration, Methodology, Investigation, Funding acquisition, Conceptualization.

Ethics statement

Participants gave informed consent before participating in the online survey where an affirmative reply was required to enter the survey. The study was explained to consumers in the online survey. They were able to withdraw from the survey at any time without giving a reason. They were financially compensated for their participation.

Funding

Funding for this research was provided by the Federal Ministry of Food and Agriculture (BMEL) based on a decision of the Parliament of the Federal Republic of Germany via the Federal Agency for Agriculture and Food (BLE) under the innovation support program (Grant number 28N1800021).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Appendix A. Appendix

Table A1 provides an overview of the three different labeling schemes and their requirements as they existed in German grocery stores during the data collection period of this study

Table A1
Overview of the three labeling schemes including examples of their requirements regarding animal welfare using the example of pigs up to 110 kg.

Label	Pictogram	Housing and Space Requirements	Activity Requirements	Feed Requirements
German organic label		Plenty of daylight, natural ventilation and outdoor access; max. 50 % of the floor area may be covered with slatted flooring. Dry, littered lying areas must be provided; minimum area ~ 75 % more space than required by law	Movement areas for mucking out and rooting as well as suitable rooting materials (e.g. straw) must be available.	100 % organic feed; at least 20 % from the own farm or from the region; genetically modified feed prohibited

(continued on next page)

Table A1 (*continued*)

Label	Pictogram	Housing and Space Requirements	Activity Requirements	Feed Requirements
Label of the animal welfare initiative		Stable housing: Minimum area 10 % more space than required by law; farms must ensure a minimum amount of daylight in the stables	The animals are provided with toys to satisfy their play instinct. In the case of pigs, the toys must be made of natural materials such as wood, sisal and natural rubber.	The pigs are constantly provided with roughage such as straw, hay or bran in addition to the regular feed supply.
Husbandry 1 "stable housing"		Stable housing: Minimum area required by law	Dry litter, which can be used for pecking, scratching and dust bathing	Feed needs to be approved or recognized by quality assurance
Husbandry 2 "stable housing plus"		Stable housing: Minimum area 10 % more space than required by law	Organic material made of changeable and consumable material such as straw; for every 150 m^2 at least one item	Feed needs to be approved or recognized by quality assurance
Husbandry 3 "outdoor access"		Stable housing with outdoor access; at least open front stable: Minimum area 40 % more space than required by law	Organic material made of changeable and consumable material such as straw; for every 150 m^2 at least two items	Feed without genetic engineering, during the entire fattening phase
Husbandry 4 "premium"		Stable housing with constant outdoor access or free range: Minimum area 100 % more space than required by law	Additional bedding in the form of straw, wood shavings, sand or peat on at least 1/3 of the barn area	Feed without genetic engineering, during the entire fattening phase; at least 20 % from the own farm or from the region

Appendix B. Appendix

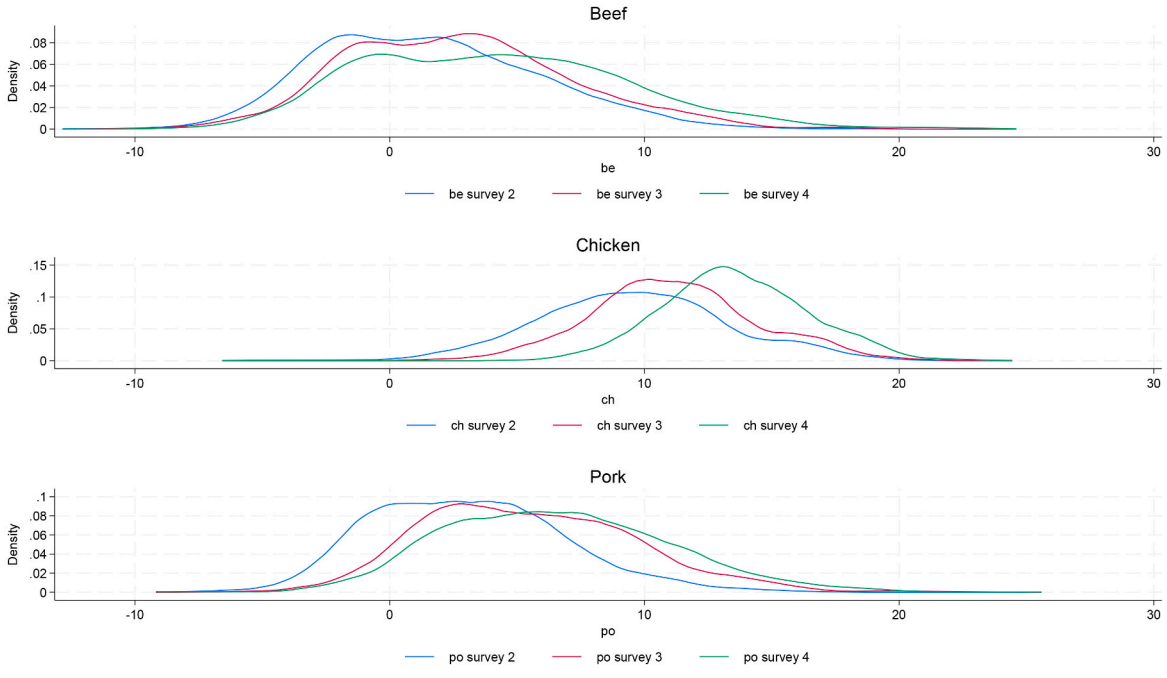

Fig. B1. Kernel density distribution for individual coefficients for the product-specific constant

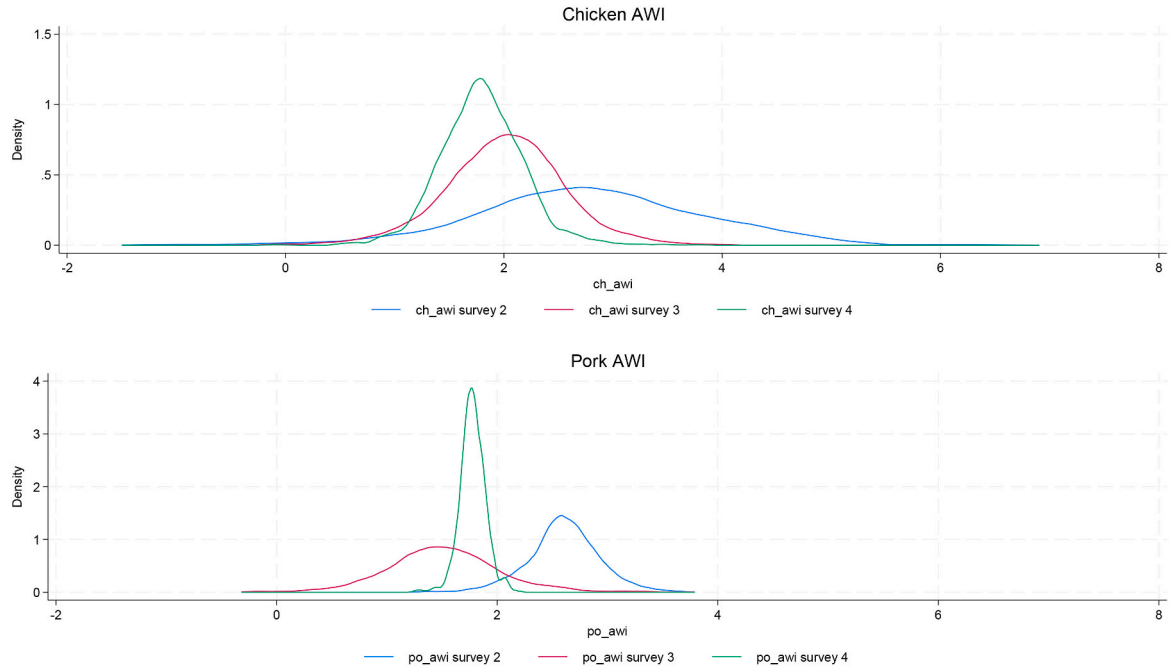

Fig. B2. Kernel density distribution for individual coefficients for the AWI label

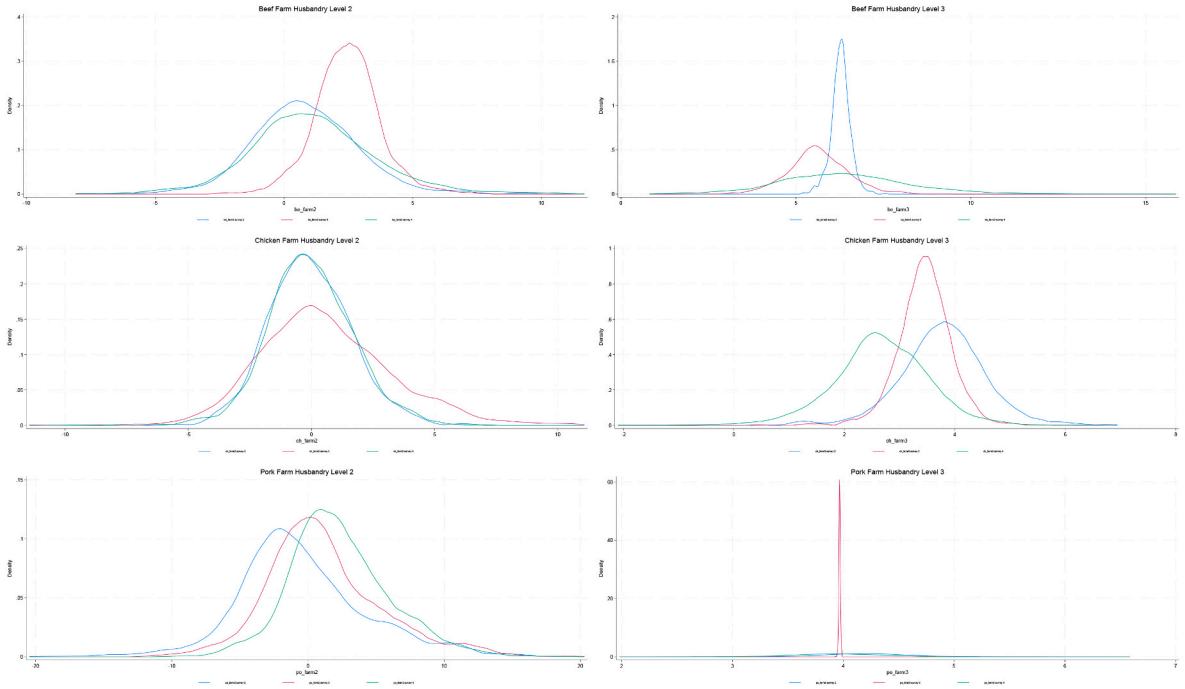

Fig. B3. Kernel density distribution for individual coefficients for the husbandry labels levels 2 and 3.

Appendix C. Appendix

Table C1
WTP per label regressed on respondents' characteristics – beef.

Variable	WTP organic	WTP AWI	WTP Husb. 2	WTP Husb. 3	WTP Husb.4
Survey 1	Base				
Survey 2	−0.272		1.628***	−0.645***	−0.430***
	(0.349)		(0.0791)	(0.0473)	(0.136)
Survey 3	−0.272		0.279***	0.0855*	−0.708***
	(0.355)		(0.0808)	(0.0484)	(0.138)
Don't know the label	Base				
Know and low trust	−1.983***		0.114	−0.0570	−0.155
	(0.297)		(0.0920)	(0.0550)	(0.157)
Know and high trust	0.0783		−0.0963	0.0437	0.335**
	(0.290)		(0.0827)	(0.0495)	(0.142)
Not female	Base				
Female	−0.168		−0.0572	−0.0361	0.157
	(0.287)		(0.0651)	(0.0389)	(0.111)
Age	−0.00542		0.00201	0.000656	−0.00168
	(0.0103)		(0.00232)	(0.00139)	(0.00398)
Income <1500	Base				
Income 1500–2599	1.368***		−0.0576	−0.0182	0.177
	(0.402)		(0.0911)	(0.0545)	(0.156)
Income 2600–3199	2.309***		0.134	0.00548	0.514***
	(0.496)		(0.112)	(0.0671)	(0.192)
Income 3200–4499	2.763***		−0.0995	0.0811	0.261
	(0.446)		(0.101)	(0.0603)	(0.173)
Income ≥4500	4.338***		−0.150	0.00891	0.943***
	(0.477)		(0.108)	(0.0645)	(0.185)
Constant	3.920***		0.794***	6.251***	10.48***
	(0.674)		(0.151)	(0.0903)	(0.258)
R-squared	0.043		0.119	0.074	0.017

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Income = Monthly net household income in Euros, N = 3700 observations.

Table C2
WTP per label regressed on respondents' characteristics – Pork.

Variable	WTP organic	WTP AWI	WTP Husb. 2	WTP Husb. 3	WTP Husb.4
Survey 1	Base				
Survey 2	−0.339	−1.090***	1.533***	−0.242***	−1.094***
	(0.244)	(0.0145)	(0.170)	(0.0139)	(0.235)
Survey 3	−0.364	−0.824***	2.751***	−0.270***	−1.399***
	(0.248)	(0.0147)	(0.173)	(0.0142)	(0.240)
Don't know the label	Base				
Know and low trust	−1.132***	−0.0233	−0.164	−0.0121	−0.150
	(0.208)	(0.0152)	(0.197)	(0.0162)	(0.273)
Know and high trust	0.394*	0.0340**	0.0204	0.00264	0.723***
	(0.203)	(0.0142)	(0.177)	(0.0145)	(0.245)
Not female	Base				
Female	−0.0239	−0.00853	−0.717***	−0.00524	0.309
	(0.201)	(0.0119)	(0.139)	(0.0114)	(0.193)
Age	−0.0141*	0.000924**	0.00306	9.29e-06	0.00282
	(0.00721)	(0.000426)	(0.00498)	(0.000408)	(0.00689)
Income <1500	Base				
Income 1500–2599	0.705**	0.00587	−0.429**	−0.00582	0.330
	(0.282)	(0.0167)	(0.195)	(0.0160)	(0.270)
Income 2600–3199	1.582***	−0.0192	−0.267	0.0123	1.171***
	(0.347)	(0.0206)	(0.240)	(0.0197)	(0.333)
Income 3200–4499	1.637***	−0.00524	−1.121***	0.00233	1.465***
	(0.312)	(0.0185)	(0.216)	(0.0177)	(0.299)
Income ≥4500	2.448***	0.00559	−1.336***	0.00791	2.076***
	(0.334)	(0.0198)	(0.231)	(0.0190)	(0.320)
Constant	0.643	2.560***	0.427	4.211***	7.053***
	(0.472)	(0.0277)	(0.323)	(0.0265)	(0.448)
R-squared	0.033	0.624	0.081	0.109	0.028

Standard errors in parentheses. *** p < 0.01, ** p < 0.05, * p < 0.1. Income = Monthly net household income in Euros, N = 3700 observations.

Data availability

Data will be made available on request.

References

Akaichi, F., Glenk, K., & Revoredo-Giha, C. (2022). Bundling food labels: What role could the labels "Organic," "Local" and "Low Fat" play in fostering the demand for animal-friendly meat. *Agribusiness, 38*(2), 349–370. https://doi.org/10.1002/agr.21733

Bernard, J. C., Duke, J. M., & Albrecht, S. E. (2019). Do labels that convey minimal, redundant, or no information affect consumer perceptions and willingness to pay? *Food Quality and Preference, 71*, 149–157. https://doi.org/10.1016/j.foodqual.2018.06.012

BLE Projektgruppe Ökolandbau. (2023). Bio-Siegel. https://www.oekolandbau.de/bio-siegel/.

BMEL. (2022). Bio-Siegel. https://www.bmel.de/DE/themen/landwirtschaft/oekologischer-landbau/bio-siegel.html.

Boaitey, A., & Minegishi, K. (2020). Who are farm animal welfare conscious consumers? *British Food Journal, 122*(12), 3779–3796. https://doi.org/10.1108/BFJ-08-2019-0634

Bolos, L. A., Lagerkvist, C. J., Edenbrant, A. K., & Nayga, R. M. (2022). Consumer preferences for visually sub-optimal food: Role of information framing and personal goals. *Resources, Conservation and Recycling, 184*, Article 106426. https://doi.org/10.1016/j.resconrec.2022.106426

Carlsson, F., Kataria, M., Lampi, E., Nyberg, E., & Sterner, T. (2022). Red, yellow, or green? Do consumers' choices of food products depend on the label design? *European Review of Agricultural Economics, 49*(5), 1005–1026. https://doi.org/10.1093/erae/jbab036

Christoph-Schulz, I., Faletar, I., Hartmann, M., Kenning, P., Luy, J., Mergenthaler, M., & von Meyer-Höfer, M. (2024). Work package 1: Monitoring of the social acceptance of livestock farming. *Journal of Consumer Protection and Food Safety, 19*(S1), 7–11. https://doi.org/10.1007/s00003-024-01485-w

Clark, B., Stewart, G. B., Panzone, L. A., Kyriazakis, I., & Frewer, L. J. (2017). Citizens, consumers and farm animal welfare: A meta-analysis of willingness-to-pay studies. *Food Policy, 68*, 112–127. https://doi.org/10.1016/j.foodpol.2017.01.006

Dusel, S., & Wieck, C. (2023). Evidence gaps hinder animal welfare progress in the European Union. *Nature Food, 4*(5), 348–349. https://doi.org/10.1038/s43016-023-00743-8

Feucht, Y., & Zander, K. (2018). Consumers' preferences for carbon labels and the underlying reasoning. A mixed methods approach in 6 European countries. *Journal of Cleaner Production, 178*, 740–748. https://doi.org/10.1016/j.jclepro.2017.12.236

Frey, U. J., & Pirscher, F. (2018). Willingness to pay and moral stance: The case of farm animal welfare in Germany. *PLoS One, 13*(8), Article e0202193. https://doi.org/10.1371/journal.pone.0205551

Gesellschaft zur Förderung des Tierwohls in der Nutztierhaltung. (2023). *https://www.haltungsform.de/*.

Gross, S., Waldrop, M. E., & Roosen, J. (2021). How does animal welfare taste? Combining sensory and choice experiments to evaluate willingness to pay for animal welfare pork. *Food Quality and Preference, 87*, Article 104055. https://doi.org/10.1016/j.foodqual.2020.104055

Grunert, K. G., Sonntag, W. I., Glanz-Chanos, V., & Forum, S. (2018). Consumer interest in environmental impact, safety, health and animal welfare aspects of modern pig production: Results of a cross-national choice experiment. *Meat Science, 137*, 123–129. https://doi.org/10.1016/j.meatsci.2017.11.022

Hasselbach, J. L., & Roosen, J. (2015). Consumer heterogeneity in the willingness to pay for local and organic food. *Journal of Food Products Marketing, 21*(6), 608–625. https://doi.org/10.1080/10454446.2014.885866

Hempel, C., & Roosen, J. (2024). Growing importance of price: Investigating food values before and during high inflation in Germany. *Agricultural Economics, 55*(6), 1026–1039. https://doi.org/10.1111/agec.12865

Hole, A. R. (2007). Fitting mixed logit models by using maximum simulated likelihood. *The Stata Journal, 7*(3), 388–401. https://doi.org/10.1177/1536867X0700700306

Initiative Tierwohl. (2019). Unkompliziert, einheitlich, nachvollziehbar: Was die vier Stufen der Lebensmittelkennzeichnung Haltungsform bedeuten. https://initiative-tierwohl.de/2019/02/08/unkompliziert-einheitlich-nachvollziehbar/.

Initiative Tierwohl. (2025). Über uns. https://initiative-tierwohl.de/initiative/ueber-uns/.

de Jonge, J., van der Lans, I. A., & van Trijp, H. C. (2015). Different shades of grey: Compromise products to encourage animal friendly consumption. *Food Quality and Preference, 45*, 87–99. https://doi.org/10.1016/j.foodqual.2015.06.001

Lagerkvist, C. J., Berthelsen, T., Sundström, K., & Johansson, H. (2014). Country of origin or EU/non-EU labelling of beef? Comparing structural reliability and validity of discrete choice experiments for measurement of consumer preferences for origin and extrinsic quality cues. *Food Quality and Preference, 34*, 50–61. https://doi.org/10.1016/j.foodqual.2013.12.009

Lagerkvist, C. J., Carlsson, F., & Viske, D. (2006). Swedish consumer preferences for animal welfare and biotech: A choice experiment. *AgBioForum, 9*, 51–58. http://hdl.handle.net/10355/96.

Lagerkvist, C. J., & Hess, S. (2011). A meta-analysis of consumer willingness to pay for farm animal welfare. *European Review of Agricultural Economics, 38*(1), 55–78. https://doi.org/10.1093/erae/jbq043

Latacz-Lohmann, U., & Schreiner, J. A. (2018). Assessing consumer and producer preferences for animal welfare using a common elicitation format. *Journal of Agricultural Economics, 70*(2), 293–315. https://doi.org/10.1111/1477-9552.12297

Lusk, J. L., & Tonsor, G. T. (2016). How meat demand elasticities vary with price, income, and product category. *Applied Economic Perspectives and Policy, 38*(4), 673–711. https://doi.org/10.1093/aepp/ppv050

Malone, T., & Lusk, J. L. (2018). A simple diagnostic measure of inattention bias in discrete choice models. *European Review of Agricultural Economics, 45*, 455–462.

Marette, S. (2014). Economic benefits coming from the absence of labels proliferation. *Journal of Agricultural and Food Industrial Organization, 12*(1), 65–73. https://doi.org/10.1515/jafio-2013-0014

McFadden, D. (1973). Conditional logit analysis of qualitative choice behavior. *Frontiers in Econometrics*, 105–142. https://escholarship.org/content/qt61s3q2xr/qt61s3q2xr.pdf.

Meas, T., Hu, W., Batte, M. T., Woods, T. A., & Ernst, S. (2015). Substitutes or complements? Consumer preference for local and organic food attributes. *American Journal of Agricultural Economics, 97*(4), 1044–1071. https://doi.org/10.1093/ajae/aau108

Mørkbak, M. R., & Olsen, S. B. (2015). A within-sample investigation of test–retest reliability in choice experiment surveys with real economic incentives. *Australian Journal of Agricultural and Resource Economics, 59*(3), 375–392. https://doi.org/10.1111/1467-8489.12067

Ngene, C. (2018). *1.2 User Manual & Reference Guide*. Sydney, Australia: ChoiceMetrics Pty Ltd. https://files.choice-metrics.com/NgeneManual.pdf.

Petersen, T., Denker, T.-L., Koppenberg, M., & Hirsch, S. (2024). Revealed preferences on meat substitute consumption and political attitudes - testing the left-right and environmental concerns framework. *Appetite, 199*, Article 107371. https://doi.org/10.1016/j.appet.2024.107371

Rigby, D., Burton, M., & Pluske, J. (2016). Preference stability and choice consistency in discrete choice experiments. *Environmental and Resource Economics, 65*, 441–461. https://doi.org/10.1007/s10640-015-9913-1

Schwickert, H. (2023). Paying for animal welfare labelling no matter what? A discrete choice experiment. *European Review of Agricultural Economics, 50*(5), 1754–1795. https://doi.org/10.1093/erae/jbad027

Sonntag, W. I., Lemken, D., Spiller, A., & Schulze, M. (2023). Welcome to the (label) jungle? Analyzing how consumers deal with intra-sustainability label trade-offs on food. *Food Quality and Preference, 104*, Article 104746. https://doi.org/10.1016/j.foodqual.2022.104746

Staudigel, M., & Trubnikov, A. (2022). High price premiums as barriers to organic meat demand? A hedonic analysis considering species, cut and retail outlet. *Australian Journal of Agricultural and Resource Economics, 66*(2), 309–334. https://doi.org/10.1111/1467-8489.12472

Thøgersen, J., & Nielsen, K. S. (2016). A better carbon footprint label. *Journal of Cleaner Production, 125*, 86–94. https://doi.org/10.1016/j.jclepro.2016.03.098

Train, K. E. (2009). *Discrete choice methods with simulation*. Cambridge University Press.

Weinrich, R., & Spiller, A. (2016a). Developing food labelling strategies: Multi-level labelling. *Journal of Cleaner Production, 137*, 1138–1148. https://doi.org/10.1016/j.jclepro.2016.07.156

Weinrich, R., & Spiller, A. (2016b). Can a multi-level label do better than a binary label for animal welfare? A PLS-analysis of consumer satisfaction. *International Food and Agribusiness Management Review, 19*(3), 1–30. https://doi.org/10.22004/ag.econ.244646

Werle, C. O., Pruski Yamim, A., Trendel, O., Roche, K., & Nadaud, P. (2022). When detailed information works better: Comparison of three-and five-color/letter front-of-package nutrition labels. *Journal of Public Policy & Marketing, 41*(2), 177–195. https://doi.org/10.1177/07439156211061289

Wilson, L., & Lusk, J. L. (2020). Consumer willingness to pay for redundant food labels. *Food Policy, 97*, Article 101938. https://doi.org/10.1016/j.foodpol.2020.101938

Yang, W., & Renwick, A. (2018). Consumer willingness to pay price premiums for credence attributes of livestock products–a meta-analysis. *Journal of Agricultural Economics, 70*(3), 618–639. https://doi.org/10.1111/1477-9552.12323

Yoo, H. I. (2020). lclogit2: An enhanced command to fit latent class conditional logit models. *The Stata Journal, 20*(2), 405–425. https://doi.org/10.1177/1536867X20931003