This version: 26 Feb 2024. Draft prepared for AES Edinburgh 2024.

Combining Geographical Indication Ham with Nutri-Score does not cause Label Fatigue

Gero Laurenz Höhn, Martijn Huysmans and Christophe Crombez

Abstract

The EU aims to propose a new harmonised nutrition label. However, the effects of colour-coded candidates such as the Nutri-Score in combination with other food labels such as Geographical Indications remain underexplored. If consumers suffer from label fatigue, the consumer valuation of combining a Geographical Indication (GI) with a better Nutri-Score would be lower than the sum of the labels separately. Therefore, we conduct a discrete choice experiment with over 800 German and Dutch respondents to quantify the willingness to pay for these labels. We find that consumers are willing to pay a premium of 72 cents for GI-labelled Parma ham and 48 cents for Nutri-Score D rather than E. We find no significant interaction term: there is no evidence of label fatigue or simplifying heuristics when combining Geographical Indications with Nutri-Score.

Keywords: Food labelling, Geographical indications, Protected Designation of Origin (PDO), Nutri-Score, Discrete Choice Experiment

JEL classifications: D12; Q13;

Gero Laurenz Höhn is Research Associate at the LICOS Centre for Institutions and Economic Performance, KU Leuven, Belgium. Martijn Huysmans is Associate Professor at the School of Economics, Universiteit Utrecht, the Netherlands. Email: m.huysmans@uu.nl for correspondence. Christophe Crombez is full professor at the Department of Management, Strategy and Innovation, KU Leuven, Belgium and Senior Research Scholar at The Europe Center, Stanford University, USA. This research was supported by KU Leuven - Long term structural funding (Methusalem funding) by the Flemish Government. The authors would like to thank seminar participants of the 9th PhD Workshop of the European Association of Agricultural Economists (EAAE) for useful comments on our research idea and plan. In particular, we want to thank Joyce Delnoij and Michael Mödl for expert advice on our experimental design and questionnaire. Furthermore, we would like to thank Veerle Siegerink and Francisco Alpízar for sharing their useful guidance on implementing discrete choice experiments in Qualtrics. Moreover, special thanks go to Daniele Curzi, Giulia Meloni, Mattia Nardotto and Jo Swinnen for comments on earlier versions. Finally, we would like to express our gratitude to all respondents that participated in our pilot and main study. All shortcomings remain our own.

1 Introduction

The European Commission aims to establish a harmonized front-of-pack label concerning products' nutritional value to guide and empower consumers in their purchase decisions (EC, 2020). Our paper's objective is to analyse the effects of a novel nutrition label on consumer evaluations of Geographical Indications (GIs). Today, consumers are confronted with more and more labels on product packages and hence, investigating how several labels interact with each other has become an important are of research for food labelling research (Drugova et al., 2020; Sonntag et al., 2023). Sonntag et al. (2023) describe this new setting for consumers as a 'label jungle'. Related label fatigue –meaning that consumers are overwhelmed by too many labels and do not take all of them equally into account – has often been addressed in the press (cf. Gunlock, 2015; von Massow, 2019). However, especially extant GI research has not properly accounted for interactions with new sustainability-related labels. Thus, we examine the interaction of a GI with labels assigning differing nutritional quality based on the Nutri-Score.

With a discrete choice experiment (DCE) we determine the willingness to pay (WTP) of German and Dutch consumers for GI-protected Parma ham and Nutri-Scores. Consumers in our sample prefer Protected Designation of Origin (PDO) Parma ham as well as a better Nutri-Score (D instead of E). The WTP of 0.72 € per 100 grams for a PDO product is higher compared to the WTP for the better Nutri-Score D, which is 0.48 € per 100 grams. We also tested for an interaction effect. If consumers suffer from label fatigue, the bonus from two positive labels (a PDO and a better Nutri-Score in our case) would be smaller than the bonus of each label separately. In other words, there would be a negative interaction term. While we estimate a negative effect suggesting a reduction of marginal utility, it is not statistically different from zero in the main regressions. Consequently, having a comparatively better Nutri-Score D does not significantly reduce the marginal utility of PDO Parma ham and vice-versa.

All in all, our main contributions are as follows. First, we quantify the relative strength of the GI and better Nutri-Score labels in affecting consumer WTP, based on a sample of more than 800 German and Dutch respondents. We find that the effect of the GI certification of PDO Parma ham, on which we focus, is stronger. Second, we investigate potential label fatigue and signs of related heuristics, such as disregarding certain labels, through an interaction effect of the PDO Parma ham with a comparatively better Nutri-Score D. In our experimental setting, the average consumer shows no strong signs of information overload or label fatigue, as they generally value both labels independently.

2 Geographical Indications and additional labelling

First and foremost, GIs are meant to certify a specific origin and production method (Moschini et al., 2008). However, the GI label got company of labels certifying other product characteristics such as nutritional quality due to an emerging trend towards additional sustainability-related labels (Sonntag et al., 2023). Gracia and de-Magistris (2010) show in their ranking experiment of various food labels that the PDO label and the nutrition facts panel were the most important labels to consumers.

2.1 Willingness to pay for GIs and Nutri-Scores

A major enabler for sustainable economic success of GIs is the realization of price premiums. Many GI hams are traded in extra- and intra-EU markets (Török & Jambor, 2016) and generally tend to achieve rather high premiums compared to other categories (AND-International, 2019). Still, consumer responses may differ in northern Member States because GI recognition and awareness remain low there (AND-International, 2020).[1] To the best of our knowledge, we contribute to the literature by conducting the first WTP analysis with German and Dutch consumers for a foreign GI product.[2]

One of the most well-known GI products is PDO Parma ham, which was analysed by various stated preference analyses. For example, Garavaglia and Mariani (2017) unravel that overall Italian consumers are willing to pay a significant premium of more than 10% for Parma ham. Van Ittersum et al. (2007) analyse, amongst others, attitudes towards Parma ham of Italian consumers. Overall, a favourable image of the PDO label significantly influences WTP of consumers (van Ittersum et al., 2007).

The Nutri-Score evaluates nutritional quality, scoring foods and classifying them into five color-coded categories A (green, best) to E (red, worst) (Julia & Hercberg, 2017).[3] Cheeses and meats, over a third of all GIs (eAmbrosia, 2023), tend to achieve higher economic value (AND-International, 2019), but score poorly on this metric, often falling into the bad categories D and E (Höhn et al., 2023b). This has been criticized by opponents like Italy and traditional

[1] For example, in 2019/20, less than a quarter of German respondents knew about the GI logos and terms and not more than a tenth of Dutch respondents (AND-International, 2020).

2 Several DCEs confirm that higher WTP for GI products also exists in newer Member States such as Hungary (Török et al., 2022) and Slovenia (Kos Skubic et al., 2018), but major northern EU markets such as Germany and the Netherlands remain underexplored. In Germany, Teuber (2011) investigates consumer expectations towards Hessian apple wine and the study of van Ittersum et al. (2007) examines, amongst others, Dutch consumer appreciation and WTP for GI cheese and potatoes. However, mentioned studies consider domestic GI products.

3 The Nutri-Score attempts to assess overall nutritional quality by summing up core macronutrients reported in the nutrition facts tables (Julia & Hercberg, 2017). Its algorithm for solid foods results in a score ranging from -15 (best) to 40 (worst). The score is then classified into one of five colour-coded categories: A (green) for -15 to -1, B (light green) for 0 to 2, C (yellow) for 3 to 10, D (orange) for 11 to 18 and E (red) for 19 to 40.

food producers, who accuse the label of unfairly penalizing their products (Fortuna et al., 2022; Borrillo, 2021; Qualivita, 2022; Roquefort, 2022).

The European Commission may propose an EU-wide label similar to the Nutri-Score (Fortuna, 2022). Germany introduced the Nutri-Score as a voluntary label in 2020 (BMEL, 2020) and in the Netherlands, during our data collection it was in a so-called 'pilot-phase' before the official introduction in 2024 (Rijksoverheid, 2023). Though not mandatory in these countries, the Nutri-Score gained popularity due to commitment of major retailers (Jumbo, 2022; REWE, 2022).[4] Thus, even producers that refuse the Nutri-Score, such as Roquefort (Roquefort, 2022), or emphasise health benefits, such as Parma or San Daniele ham (Parma, 2023b; San Daniele, 2022), may still bear a red E in main EU export markets (e.g. Parma, 2023a). Thus, it is still highly relevant to examine consumer preferences for GIs with bad Nutri-Scores.

Previous literature on the Nutri-Score focused on consumer understanding and appeal, highlighting the Nutri-Score as intuitive and easily recognisable (Becker et al., 2015; Egnell et al., 2020; Hagmann & Siegrist, 2020). However, limited research exists that investigates the WTP for Nutri-Scores, which is important to quantify as higher WTP for better Nutri-Scores represents another incentive for producers to reformulate. In the study of Gassler et al. (2022), a sample of German consumers is also willing to pay a mark-up for a Nutri-Score B and A on yogurts. These findings are confirmed in two DCEs considering whole milk and chicken breast in the study of Sonntag et al. (2023). German consumers have a higher WTP for the Nutri-Score B and a negative WTP for the worse Nutri-Score D (Sonntag et al., 2023).[5] Consequently, consumers tend to prefer products with better Nutri-Scores and are willing to pay a considerable premium. However, little is yet known how comparatively similar Nutri-Scores are valued by consumers. Thus, we quantify the difference in WTP between Nutri-Scores assigning low(er) nutritional value, namely categories D and E.

Moreover, the interplay of the Nutri-Score with other quality-related labels such as GIs remains underinvestigated. In an exploratory study, Stiletto and Trestini (2022) provide empirical evidence that the marginal WTP for PDO cheeses is higher than for the Nutri-Score.[6] Nonetheless, they found that overall consumers had a higher WTP for the 'unhealthy' Nutri-

4 For example, the German retailer REWE decided to commit to the Nutri-Score label, especially regarding their store brands (REWE, 2022) and the Dutch retailer chain Jumbo even set the goal to increase the number of 'healthy' private label products with A and B Nutri-Scores by 2% by 2023 (Jumbo, 2022).

5 German consumers have a higher WTP for the Nutri-Score B, with an increase of about 0.53 € per 100 grams of chicken breast (N=482) and 0.30 € per litre of whole milk (N=503) (Sonntag et al., 2023).

6 Stiletto and Trestini (2022) investigate WTP of Italian respondents for GIs and showing a Nutri-Score D in the case of the well-known PDO Asiago (N=300) and lesser-known (N=300) PDO Casatella Trevigiana.

Score D, which was the only Nutri-Score category considered (Stiletto & Trestini, 2022). However, just comparing a Nutri-Score D to no Nutri-Score leads to two concerns. If the Nutri-Score becomes mandatory, the scenario without the Nutri-Score is no longer relevant. Although the countries endorsing the Nutri-Score did not make it yet mandatory de jure, labels like the Nutri-Score may become de facto mandatory due to retailers' commitment and bargaining power (Lemken et al., 2021). Secondly, while Stiletto and Trestini (2022) assert that the Nutri-Score D is a bad score, consumers are not given an explanation or comparison to products with different Nutri-Scores in the category of cheese. Therefore, the implicit comparison may be the worse Nutri-Score E rather than C or even more favourable scores. The latter may explain why some Italian respondents were willing to pay a premium for the Nutri-Score D in their study. Due to these concerns, we consider varying Nutri-Scores on a PDO-labelled product which is sold in countries that already endorse the Nutri-Score.

2.2 Behavioural responses to multiple food labels: Fatigue and heuristics

Stiletto and Trestini (2022) also interact the PDO certifications with a Nutri-Score D, yet their results remain inconclusive overall. Given the general proliferation of sustainability-related labels on foods, recent studies started exploring interaction effects of various food labels (Drugova et al., 2020; Gassler et al., 2022; Sonntag et al., 2023). We aim to contribute to this emerging and topical literature by interacting a GI certification with a (better) Nutri-Score.

Combining different labels, e.g. certifying organic production and animal welfare, can lead to a reduction in marginal utility. In other words, the WTP for the combination is less than the WTP of each individual label together (Gerini et al., 2016). The reduction of marginal utility has been described in the media as 'label fatigue', i.e. that more labels do not necessarily generate more value and are potentially ignored (Gunlock, 2015; von Massow, 2019).

This label fatigue phenomenon can be tied back to fundamental concepts in behavioural economics: choice fatigue and information overload. An overabundance of product options to choose from can in certain cases lead to dissatisfaction and suboptimal decisions of consumers due to an information overload that is difficult to process (Chernev et al., 2015; Scheibehenne et al., 2010). Searching for the preferred product can be a straining process for consumers (Carlin & Ederer, 2018). Contrary to the classical notion of consistently rational consumers, research in behavioural economics suggests that decisions are not only based on maximizing utility according to preferences, but are also influenced by systematic biases, heuristics and context (Reisch & Zhao, 2017).

Diving deeper, Steenkamp (1990) already posited that given the multitude of quality cues in the marketing environment, consumers, bounded by cognitive constraints, can only consider a fraction of them. As such, they often base their perceptions on a limited set of these indicators. Rappoport et al. (1993) emphasise in their food cognition model that the choice to consume a specific food hinges on its rating across three distinct criteria: pleasure, health and convenience. However, consumers may apply simplifying heuristics to ease their choice due to label fatigue. Consider so-called 'vice' foods like ham, ice cream or cheese. These are primarily consumed for their pleasurable taste (van Doorn & Verhoef, 2011).[7] As a result, consumers tend to overlook e.g. the nutritional value of vice products (Drugova et al., 2020). Bettman (1979) introduced the lexicographic decision rule, a straightforward heuristic suggesting individuals make choices based on a single predominant reason. Specifically, they opt for the alternative with the highest rating on their most valued attribute, such as selecting the healthiest food. Supporting this notion, Scheibehenne et al. (2007) highlight that the lexicographic heuristic, where consumers focus primarily on their top concerns, holds potential in predicting food choices.

Building on the existing research, our study aims to enrich the broader literature on consumer behaviour and choice heuristics by empirically exploring potential fatigue arising from the presence of multiple food labels on hams. Notably, while the GI label leans towards signalling tradition, craftsmanship, and gustatory quality and is thus more related to hedonistic aspects, the Nutri-Score directly addresses health and nutritional quality. Our experiment intends to examine if consumers value both criteria individually or if they tend to ignore one in favour of the other, suggesting label fatigue and the use of lexicographic heuristics in the case of ham.

3 Method and data

3.1 Hypotheses

On the basis of the aforementioned literature we define our main hypotheses to test. To begin with, we expect a higher WTP for the PDO Parma ham certification than for the comparatively better Nutri-Score D. The underlying reason for this expectation is that nutritional quality just represents one credence attribute of the product. In contrast, GIs such as Parma ham follow

[7] In general, vice goods can include tobacco products and alcoholic beverages. In our context, vice goods refer to savoury and flavourful, but (rather) unhealthy foods such as cured meats. The latter should be consumed more moderately due to high salt and fat contents as well as potential carcinogenic risks (WHO, 2020; WHO, 2015). Thus, vice goods are described as vices because they satisfy instant desires, such as pleasure from sensory quality, rather than long-term goals, such as avoiding weight gain and health issues (Doorn & Verhoef, 2011).

product specifications that not only certify origin, but possibly also other relevant experience, credence or Potemkin attributes such as taste due to e.g. minimum maturation time or authenticity by stipulating more traditional production techniques (Huysmans et al., 2022). Moreover, consumers tend to care less about sustainability-related labels such as organic production or nutritional quality with regard to vice products. 'Vice' goods are goods that consumers consume mainly for pleasure (e.g. they taste good) rather than for healthiness (Drugova et al., 2020; van Doorn & Verhoef, 2011). Therefore, our first hypothesis reads:

(H1 – Vice Good) Consumers have a stronger preference for PDO Parma ham than for a better Nutri-Score

Second, our paper examines whether there is an additional benefit of combining PDO Parma ham with a better Nutri-Score D. If consumers suffer from label fatigue, they could show signs of lexicographic heuristics and may not value both labels independently. Thus, the WTP for the combination could be less than the sum of the individual WTP values. Consequently, our second hypothesis is:

(H2 – Label Fatigue) There is a negative interaction in the WTP for a PDO Parma ham with a better Nutri-Score

3.2 Data collection and experimental design

3.1.1 Focus product and sample

For our DCE we focus on Parma ham as it is representative of popular GIs sold across the EU. Moreover, studies mentioned earlier highlight price premiums, but none of them considered nutrition labels. Most importantly, Parma ham and raw ham in general do not fall into one Nutri-Score category, but can receive either D (orange) or E (red). Hence, less healthy options are available, which we double-checked in stores and online databases such as Open Food Facts.[8]

Instead of focusing on the Italian home market of Parma ham, we recruit respondents in Germany and the Netherlands. In contrast to Italy, the Nutri-Score is present and publicly endorsed in both countries. Furthermore, Germany and the Netherlands represent main intra-EU export markets of Parma ham (Parma, 2023a) and Italian food GIs in general (Qualivita, 2023). German and Dutch consumers are known to be price sensitive (Retailtrends, 2013; Statista, 2013). They also have low GI awareness (AND-International, 2020). These are different circumstances compared with the Italian home market. Therefore, it is important to

[8] See https://world.openfoodfacts.org/.

quantify WTP and understand preferences of these consumers as well. To the best of our knowledge, we contribute to the literature by conducting the first WTP analysis with German and Dutch consumers for a foreign GI product.

We administered the final study in April and May 2023 via the service provider Prolific with gender-balanced samples of German and Dutch consumers.[9] The study took about seven minutes to complete and respondents were monetarily rewarded with 1.25 £, i.e. about 11 £ per hour. We pre-screened respondents based on their diet, i.e. we did not consider respondents that follow a vegetarian, pescatarian or vegan diet and hence, do not consume meat. In total, we received 980 completed responses to our questionnaire. However, we only consider a sample of 815 valid responses because we removed respondents that took less than two minutes to complete (41) and that failed our attention check (94).[10] In addition, we removed 30 respondents from our sample because they neither purchased ham in the past six months nor did they consume ham in the past three months. Consequently, we focus on current consumers of ham who are familiar with the product shown in the choice situations.

The Appendix gives more information on the sample and how it is representative of the general population.

3.1.2 Experimental design

In a DCE, consumers choose repeatedly between different hypothetical options of a product according to their preferences for defined product characteristics (Bliemer & Rose, forthcoming). In our experiment, we consider three different product characteristics with differing levels. The first characteristic is '*Geographical Indication (GI)*' which has two levels. Respondents encounter either PDO Parma ham or generic raw ham without a GI. The second characteristic is '*Nutri-Score (NS)*'. The shown ham has either a Nutri-Score of D (orange) or E (red) because no better score would be possible for a raw ham based on the Nutri-Score metric. To determine WTP and test our hypotheses, the third attribute is '*Price / 100 grams*', which has three levels: 3€/100gr, 4€/100gr and 5€/100gr. The price levels were determined based on current in-store and online store prices in major German and Dutch retailers. We rounded prices to decrease cognitive strain (Bliemer & Rose, forthcoming) and chose the price per 100 grams because 100 grams also represents a standard package size for raw hams sold in supermarkets.

[9] Prolific (https://www.prolific.co/) is a platform tailored for researchers, offering online participant recruitment with a diverse pool of study participants from various countries. It upholds good standards in recruitment, ensuring participants are aware they are being engaged for research purposes (Palan & Schitter, 2018).

[10] Separate from the choice tasks, we asked respondents to rank the importance of the attributes. Our attention check excludes respondents who chose the same attribute as least and most important.

Table 1 Product characteristics with given explanation and levels

Characteristic	Explanation (English translation)	Levels
Geographical Indication *(GI)*	Compared to *other raw hams*, traditional raw hams with a geographical indication such as *Parma ham* must be produced in a certain area (e.g. hills around Parma in Italy). In the EU, the label that stipulates a specific origin and production requirements is called *Protected Designation of Origin* or *PDO*:  In the survey, you will encounter ***PDO Parma hams*** and ***raw hams***. The latter are not protected by a PDO label and hence, not limited in terms of origin and production.	I. Parma ham (PDO) II. Raw ham (generic)
Nutri-Score *(NS)*	The *Nutri-Score* intends to measure *the overall nutritional value* of a product and hence, to show *how healthy* the consumption is. Products are assigned to one of five colour-coded categories. Products bearing an A (green) are most likely to contribute to a healthy diet, while products with an E (red) are least likely:  Raw hams normally fall in the categories of ***D (orange)*** or ***E (red)***. In the survey, you will encounter only these two categories.	I. D (orange) II. E (red)
Price / 100 grams *(Price)*	100 grams represent a regular package size of pre-sliced hams in supermarkets. Thus, *prices are given in euro per 100 grams*: **€ / 100gr** In the survey, prices will range from ***3€ / 100gr*** over ***4€ / 100gr*** to ***5€ / 100gr***.	I. 3 €/100gr (low) II. 4 €/100gr (medium) III. 5 €/100gr (high)

Following related DCE studies, all characteristics were briefly explained to respondents (see Table 1) so that they know the basic meanings. These explanations avoid confusion among respondents and ensure that characteristics are not misunderstood or ignored due to a respondent's possible lack of knowledge about a characteristic (cf. De Bauw et al., 2022; Mazzocchi et al., 2022; Pérez y Pérez et al., 2020).

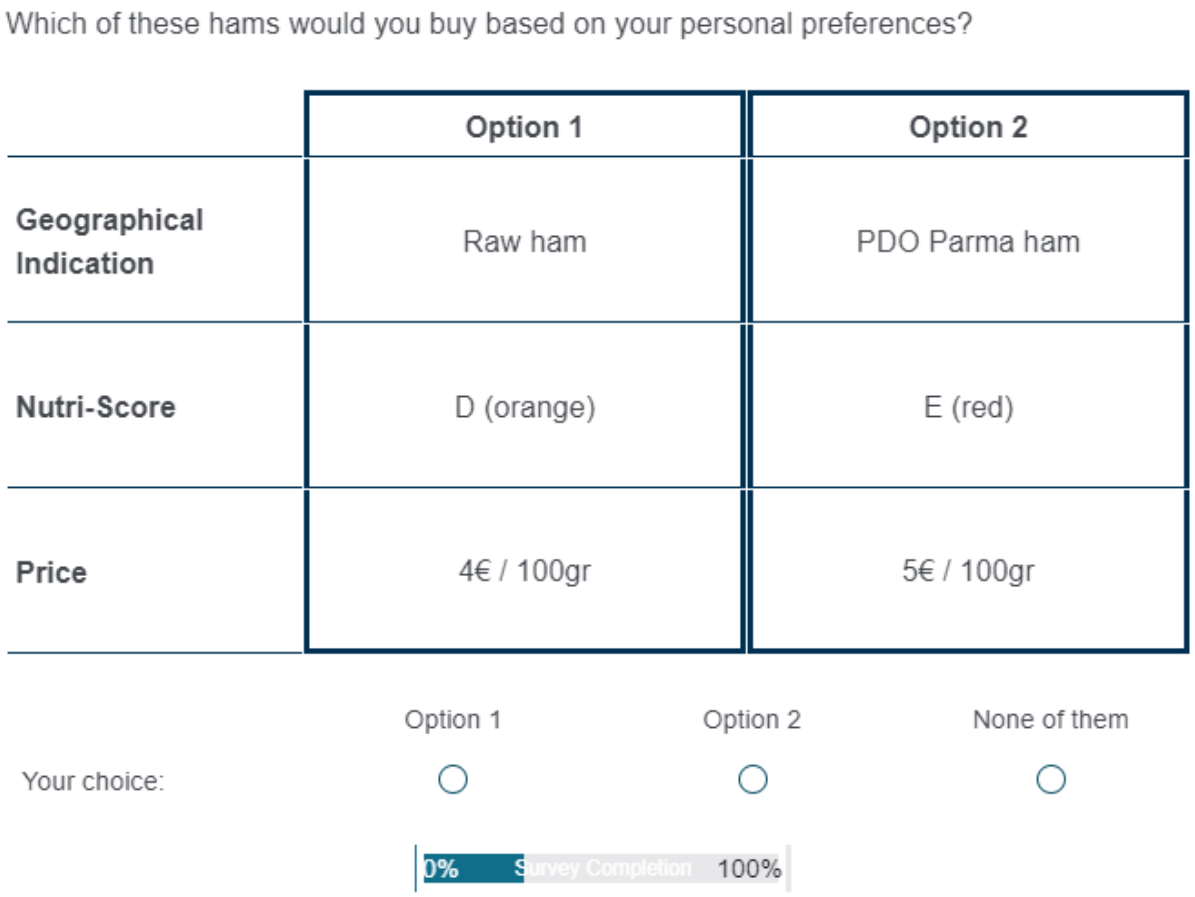

Figure 1 Exemplary choice set shown to our respondents (English translation).

Respondents had to choose their preferred ham out of two options differing with regard to the considered characteristics. We use a fractional factorial design because in total we have 66 possible comparisons in our full factorial design which is too much to show to each respondent. Based on the 66 possible comparisons we created a random design which allowed us to exclude undesirable dominant options (Bliemer & Rose, forthcoming). In our case, a dominant option would represent e.g. a ham that is GI-certified, healthier and cheaper compared to the other option. Our final random design includes 18 different choice situations that were

randomly assigned into two blocks to reduce cognitive burden.[11] Hence, each respondent had to decide in 9 choice situations. Respondents were randomly assigned to each block via Qualtrics accounting for block balance.[12]

As we follow a stated preference approach, we also showed respondents a cheap-talk script to reduce hypothetical bias (Menapace & Raffaelli, 2020; Penn & Hu, 2018).[13] In addition, we added a 'no-purchase' option to make the choice generally more realistic (Bliemer & Rose, forthcoming). Consequently, respondents could either choose 'Option 1', 'Option 2' or 'None of them'. We first created an English version of our DCE questionnaire in Qualtrics and conducted a pilot study (N=38).[14] See Figure 1 for an exemplary choice situation shown to respondents. The DCE design was tested successfully and hence, not changed but we added and reformulated some explanations and questions concerning consumption behaviour and socio-demographics based on feedback from respondents and two fellow researchers. Finally, the questionnaire was translated in German and Dutch.[15]

3.3 Econometric model

Preferences of respondent *i* for choosing option *k* in choice situation *n* based on certain characteristics can be modelled according to the well-established random-utility theory (Lancaster, 1966; McFadden, 1974; Train, 2009):

$$U_{ikn} = \beta_1 Price_{ikn} + \beta_2 GI_{ikn} + \beta_3 NS_{ikn} + \beta_4 GI * NS_{ikn} + \varepsilon_{ikn}$$

$U = consumer\ utility, i = respondent, k = choice\ option, n = choice\ situation$

β_1 represents the coefficient of *Price*, which is defined as a continuous variable. β_2 is the coefficient of the *GI* dummy, which takes the value of 1 if it is PDO Parma ham. β_3 is the coefficient of the *NS* dummy, which takes the value of 1 if the product has the better Nutri-

[11] The random design also exhibits desirable attribute level balance, i.e. that there is not a certain level of a characteristics predominantly shown (Bliemer & Rose, forthcoming). For instance, half of the options have a Nutri-Score D, while the other half have E.

[12] The Qualtrics software ensures that an equal number of participants is allocated to each block. This practice prevents potential bias that could arise from an uneven distribution of participants across the blocks.

[13] A cheap-talk script is a method often employed in discrete choice experiments to reduce hypothetical bias in estimating WTP. This technique involves informing respondents of the common tendency to overstate WTP in hypothetical scenarios and urging them to consider this when expressing their preferences (Cummings & Taylor, 1999).

[14] Qualtrics is a widely-used online survey platform that allows researchers to design, distribute, and analyse complex online surveys and collect data for various research purposes (see https://www.qualtrics.com/). The link to the survey was shared with academic colleagues and peers in December 2022 to simulate the respondent experience and gather initial feedback.

[15] Our final DCE design and questionnaire was scientifically, GDPR and ethically approved by the Social and Societal Ethics Committee (SMEC) of KU Leuven with reference number 'G-2022-5603-R2(MAR)' in February 2023.

Score D (orange). β_4 is the coefficient of the interaction effect of *GI* and *NS*. Finally, ε_{ikn} represents the error term.

In the past, the conditional logit model (CLM) represented the main tool for the analysis of DCE data. However, the CLM makes the often unrealistic assumption of preference homogeneity among consumers. In order to relax this strict assumption, we use a mixed logit model (MXL) for our data analysis to account for preference heterogeneity among consumers (Train, 2009). Due to the fact that the MXL accounts for preference heterogeneity, it emerged as a new standard in DCE analysis (Caputo & Scarpa, 2022; Lizin et al., 2022).

As in the case of the CLM, the dependent variable of our MXL is binary and takes the value of 1 if an option was chosen and 0 if otherwise. However, for the MXL we have to define the remaining parameters as either random or fixed. We assume that consumer preferences differ regarding GIs and Nutri-Scores. Hence, we define the coefficients of the dummies *GI* and *NS* as random based on a normal distribution. Using such random coefficients allows for the fact that some consumers in our sample favour the characteristic and others do not. Accordingly, we also estimate the coefficient of the interaction term *GI*NS* as random. Moreover, we introduce the continuous variable of *Price* as a fixed parameter, which is common practice due to the fact that rational consumers should generally prefer lower prices. Also, we include the alternative-specific constant (ASC) *No-Purchase* for the option 'None of them'. *No-Purchase* is also random as consumers may differ in their preference regarding the choice to 'opt out'. We run all our MXL specifications with 500 Halton draws.

In the end, we determine the marginal WTP for PDO Parma ham and a Nutri-Score D based on the negative ratio of the coefficient of *GI* and *NS* respectively to the coefficient of *Price.* For the interaction, we use the negative ratio of the sum of the coefficients *GI*, *NS* and *GI*NS* to the coefficient of *Price* (cf. Drugova et al., 2020). The marginal WTP represents the marginal amount in € per 100 grams that a consumer is willing to pay for a change in the corresponding product characteristic, i.e. a PDO Parma ham and/or having a better Nutri-Score D.

4 Results

The regression results are presented in Table 2. In Model 1, we only include the main effects, all of which are statistically significant at the 1% level. German and Dutch consumers in our sample clearly prefer PDO Parma ham and a better Nutri-Score D. Generally, our respondents expectedly prefer lower prices because the coefficient of *Price* is negative. The realistic nature of the choice options is also underscored by the negative coefficient for the *No-purchase* option because respondents clearly prefer to make a purchase decision.

Table 2 Results of mixed logit regressions

	Model (1)	Model (2)
Price	-3.14***	-3.16***
	(0.09)	(0.09)
No-purchase (ASC)	-12.36***	-12.35***
	(0.34)	(0.35)
PDO Parma ham (GI)	2.22***	2.27***
(base level: generic)	(0.13)	(0.14)
Nutri-Score D (NS)	1.45***	1.51***
(base level: E)	(0.12)	(0.13)
*GI*NS*		-0.23
(interaction)		(0.14)
Summary Statistics		
N	22,005	22,005
Log-likelihood	-5080.31	-5078.09
AIC	10174.61	10174.18

Notes: *p < 0.1, **p < 0.05 and ***p < 0.01. Standard errors are shown in parentheses.

In Model 2, the main effects remain highly significant and similar in magnitude once we add the interaction effect which is negative, but not significant. Thus, this finding does neither indicate strong labelling fatigue. The marginal utility for each individual attribute is not significantly decreased (or increased) if both labels are combined.

Based on our main specification of Model 2 we calculate the marginal WTP in € per 100 grams for each of the considered product characteristics (see Figure 2). The marginal WTP for the PDO ham is with 72 cents higher than the 48 cents for the better Nutri-Score D, in keeping with *H1 (Vice Good)*. Adding up the main coefficients with the interaction shows that the sum (1.13 €) is somewhat smaller than the addition of the two main effects only (1.20 €). In terms of the sign of the coefficient, the slight negative interaction of 7 cents is consistent with *H2 (Label Fatigue)*. However, keep in mind that the interaction *GI*NS* is not significant in Model 2, so we cannot reject the null hypothesis.

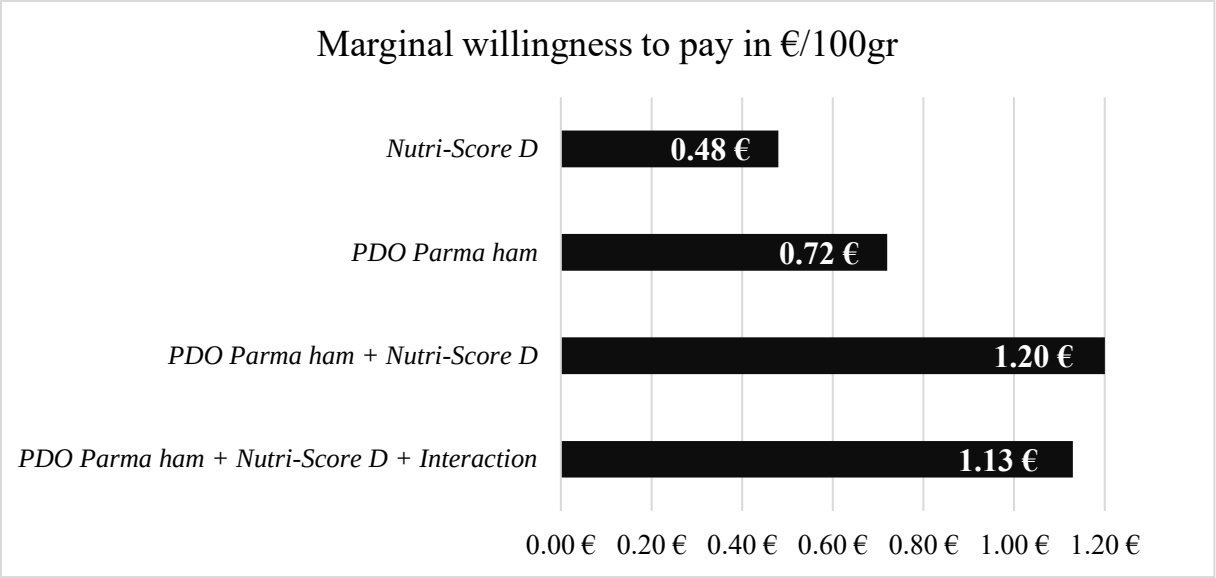

Figure 2 Marginal WTP per product characteristic and interaction thereof (Model 2).

To sum up, our findings support *H1 (Vice Good)* but not *H2 (Label Fatigue)*. First, consumers have a larger preference for PDO Parma ham than the Nutri-Score D. Second, there is no significant negative interaction when combining both PDO-labelled ham and a better Nutri-Score.

As Appendix 2 shows, there is significant heterogeneity across consumers. The standard deviations or our random coefficients show that some respondents value the attributes more than others. For instance, consumers with prior knowledge of the PDO label have a stronger preference for Parma ham and a weaker preference for the Nutri-Score D. This is consistent with the general spirit of *H1 (Vice Good)*: consumers who care about food quality labels appear to be more focused on the taste and gustatory quality of a product like ham, rather than on health.

5 Discussion and limitations

To begin with, we confirm previous studies that highlight higher WTP for GIs and Nutri-Scores. Although the sample is not fully representative of the German and Dutch populations, it still represents major societal and consumer groups. In our study, the WTP for the PDO ham is considerably higher than for the Nutri-Score D. Hence, consumers are willing to pay more for a PDO certification than for a comparatively better Nutri-Score, at least in the case of the well-known Parma ham.

The latter is also a limitation of our study as we focus on a single GI product. PDO Parma ham represents one of the most widely spread GIs. In contrast, less well-known GI hams such as Croatian ham from Dalmatia or Belgian ham from the Ardennes may show lower

premiums compared to the famous Parma ham (Leufkens, 2018). In addition, Parma ham represents a PDO which comes along with typically stricter production rules and higher price premiums (AND-International, 2019; Deselnicu et al., 2013). PGI hams, such as Black Forest ham from Germany or Tiroler ham from Austria, tend to show lower premiums (Höhn et al., 2023a). Thus, our results may apply more to well-known and advertised PDOs and should be interpreted with caution regarding the WTP for GI labels in general.

By contrast, our sample of younger Dutch and German consumers may underestimate WTP. Dutch and German consumers are known to be price-sensitive, which mirrors itself in the strong stance and growing market share of low-price discounters (Konrad, 2023; Retailtrends, 2022). This price sensitivity is likely to have contributed to an underestimation of WTP for a GI product and Nutri-Score, particularly considering our rather young sample. While the potential cases of overestimation and underestimation in our experiment might partially offset each other, the overall impact remains uncertain due to the complex interplay of various demographic and product-related influences on WTP. Further research with a more diverse and fully representative sample could provide more detailed insights into the extent to which these effects could actually offset each other.

In spite of these limitations, our respondents generally prefer hams with a better Nutri-Score, also for the GI ham. Even if some consumers appear to treat Parma ham as a vice product, which is mainly consumed for hedonistic reasons and not for nutritional benefits, the majority are willing to pay more for a Parma ham and a comparatively better Nutri-Score. This suggests a relative valuation where consumers appreciate an improvement, albeit still within an unhealthy range as neither a Nutri-Score D nor E is considered inherently healthy. GI cheeses and prepared meats typically fare worse in Nutri-Score assessments compared to their generic counterparts. Thus, GI producers might find it beneficial to explore healthier alternatives. To appeal to a more health-conscious demographic, introducing GIs with reduced salt and fat content could be advantageous, especially in products like ham where such reductions can enhance the Nutri-Score (Höhn et al., 2023b). In fact, PDO San Daniele and Parma ham amended their specifications regarding minimum sodium levels.[16] Beyond the scope of GIs, from a broader public health perspective, curbing excessive salt intake has intrinsic benefits (ANSES, 2023), even if it does not elevate the Nutri-Score category.

The negative but insignificant interaction effect of PDO Parma ham with the better Nutri-Score D aligns well with findings from previous studies. Although multiple health and

[16] See eAmbrosia database: https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/geographical-indications-register/ (status April 2023).

nutrition labels are positively valued individually, the study by Barreiro-Hurle et al. (2010) suggests that their combination does not result in additional benefits. Drugova et al. (2020) expand beyond focusing on health-related labels. In their DCE experiment with US respondents, they demonstrate that when low sugar content is combined with an organic label on cookies there is a lower WTP compared to the sum of the individual attributes. Sonntag et al. (2023) provide further evidence in a DCE considering several labels with varying levels including 'bad' and 'good' Nutri-Scores in the cases of chicken breast and whole milk. Regarding two-way interactions of positive attributes, such as organic production, lower climate impact, better Nutri-Scores and animal welfare, they find partial evidence for a reduction in marginal utility. Most of these interaction effects have indeed a negative sign, but are also insignificant. Our study echoes these findings in the case of ham with different labels because overall, our respondents do not receive significantly lower utility from the combination of labels but tend to value both independently.

Consequently, our results do not indicate strong label fatigue or lexicographic heuristics. Consumers seem to be able to cope well with two labels at the same time. Gracia and de-Magistris (2010) show in their ranking experiment that the PDO label and the nutrition facts panel were the most important labels to consumers. In our experiment, we indeed find higher WTP for PDO Parma ham and a better Nutri-Score. In contrast, Grunert and Aachman (2016) mention that while GIs are also often associated with distinctiveness and quality, their relevance regarding the decision-making of consumers depends on other cues and is thus, likely low. In our controlled setting, which intentionally limits products to three characteristics to isolate their effects, the GI (specifically PDO Parma ham) emerged as a significant cue, both in comparison to the Nutri-Score D and in terms of WTP relative to average prices.

However, in reality, the Nutri-Score is not the only label that accompanies GI labels on product packaging. With an increasing number of sustainability-related labels, there is potential for label fatigue and a shift towards lexicographic heuristics, resulting in diminished attention to the GI. For example, the Eco-Score made headlines as a new label to asses environmental performance. In a Belgian experiment, consumers even found Eco-Scores more important than organic production in the context of vegetable choices (De Bauw et al., 2022). Moreover, recent animal welfare labels such as *Haltungsform* or *Better Leven* in Germany and the Netherlands may influence consumer choices of GI products more than the Nutri-Score. Therefore, future research on label fatigue should delve into the interaction effects of a broader array of labels, such as Eco-Scores and animal welfare labels, on preferences for GIs. Controlled field experiments, revealed preference data or eye-tracking studies, could shed light on the real-world

decision-making process of consumers when faced with multiple labels. Moreover, policymakers and producers should carefully consider new labels on food products. Weighing the cognitive load of multiple labels and potentially streamlining or simplifying the information presented might be pivotal, as label proliferation could decrease consumers' utility.

6 Conclusion

In our study, we quantify jointly the WTP of German and Dutch consumers for raw hams with GI certification and comparatively better Nutri-Scores. Through an original discrete choice experiment, we find that consumers in our sample prefer PDO Parma ham over generic ham and the Nutri-Score D over E. Our results align with prior findings that highlight higher WTP for GIs and better nutritional quality based on the Nutri-Score. In addition, our findings also clearly show that respondents are willing to pay considerably more for PDO Parma ham (72 cents) than the better Nutri-Score D (48 cents). Overall, combining PDO Parma ham with a better Nutri-Score D neither increases nor significantly decreases the marginal utility. Consumers in our study, on average, appear to value and cope well with both labels, showing no evident signs of label fatigue or reliance on simplifying heuristics.

On the one hand, it must be highlighted that Parma ham is a prominent GI product with a strong brand. Hence, the estimated premium is likely to be higher than for other GI hams or the PDO label in general. On the other hand, our younger sample of Dutch and German respondents is likely to underestimate the premium for PDO Parma ham and the better Nutri-Score D. In general, our mixed logit regressions also underline that there is considerable preference heterogeneity among consumers. Partially, this can be explained by age, nationality, gender and prior PDO knowledge. Those who were familiar with the PDO label before tend to value the PDO Parma ham more and the Nutri-Score less. This meshes well with findings about 'vice' foods, which are mostly consumed out of hedonistic and not health-related motives (van Doorn & Verhoef, 2011).

Our study offers several business and policy implications. First, GI producers might consider making product specifications more flexible, such that better Nutri-Scores could be achieved within them (FAO, 2021). That would allow them to combine a GI premium with a premium for a better Nutri-Score. Second, with growing labelling initiatives highlighting various sustainability aspects separately, a more streamlined approach from policymakers may be essential. Overloading consumers with multiple labels can lead to confusion, potentially driving them to rely on heuristics instead. One possible direction could be to introduce more encompassing labels that merge related sustainability concerns. Such an approach would

minimise the need for numerous certifications, possibly striking a balance between informative clarity and consumer-friendly simplicity. Other options are to establish elevated baseline standards or to address the excess of unhealthy ingredients with reduction targets and potentially through harmonised taxation. Such initiatives would refrain from using more labels.

All in all, we hope that our study could shed some light on the WTP for GIs and better Nutri-Scores and the empirical extent of label fatigue. Our findings suggest limited label fatigue or reliance on heuristics in a dual-label scenario. However, future research should probe the effects of incorporating more than two labels evaluating their interplay on preferences and WTP across diverse food categories. In that respect, more should be invested in real-world, revealed preference scenarios. As sustainability-related labels continue to gain momentum, there is a potential risk of overwhelming the consumer. Addressing this challenge is vital, and the highlighted avenues for future research may assist in shaping and navigating the evolving food label landscape.

7 References

AND-International. (2019). *Economic value of EU quality schemes, geographical indications and traditional specialities guaranteed*. Publications Office of the European Union. https://doi.org/10.2762/396490

AND-International. (2020). *Evaluation support study on Geographical Indications and Traditional Specialities Guaranteed protected in the EU*. Publications Office of the European Union. https://data.europa.eu/doi/10.2762/891024

Barreiro-Hurle, J., Gracia, A., & de-Magistris, T. (2010). The effects of multiple health and nutrition labels on consumer food choices. *Journal of Agricultural Economics*, 61(2), 426–443. doi: 10.1111/j.1477-9552.2010.00247.x

Becker, M. W., Bello, N. M., Sundar, R. P., Peltier, C., & Bix, L. (2015). Front of pack labels enhance attention to nutrition information in novel and commercial brands. *Food Policy*, 56, 76–86. https://doi.org/10.1016/j.foodpol.2015.08.001

Bettman, J. R. (1979). *An information processing theory of consumer choice*. Addison-Wesley.

Bliemer, M. C. J., & Rose, J. M. (forthcoming). Designing and conducting stated choice experiments. In S. Hess, & A. Daly (Eds.), *The handbook of choice modelling*. Edward Elgar Publishing.

Borrillo, M. (2021, November 15). Nutriscore, patatine fritte più salutari dell'olio. Patuanelli: Italia contraria. *Corriere Della Sera*. https://www.corriere.it/economia/consumi/21_novembre_15/nutriscore-patatine-fritte-piu-salutari-dell-olio-patuanelli-italia-contraria-743ece18-4605-11ec-9a24-28e7c2e627b2.shtml

Bundesministerium für Ernährung und Landwirtschaft (BMEL). (2020). *Nutri-Score*. https://www.bmel.de/DE/themen/ernaehrung/lebensmittel-kennzeichnung/freiwillige-angaben-und-label/nutri-score/nutri-score_node.html#:~:text=Mehr-,Hilfestellung%20f%C3%BCr%20Unternehmen%20%E2%80%93%20Einf%C3%BChrung%20des%20Nutri%2DScore,des%20Lebensmittelkennzeichens%20in%20Deutschland%20m%C3%B6glich.

Caputo, V., & Scarpa, R. (2022). Methodological Advances in Food Choice Experiments and Modeling: Current Practices, Challenges, and Future Research Directions. *Annual Review of Resource Economics*, 14(1), 63–90. https://doi.org/10.1146/annurev-resource-111820-023242

Carlin, B. I., & Ederer, F. (2019). Search Fatigue. *Review of Industrial Organization*, 54, 485–508. https://doi.org/10.1007/s11151-018-9657-5

Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, 25(2), 333–358. http://dx.doi.org/10.1016/j.jcps.2014.08.002

Consorzio del Prosciutto di Parma (Parma). (2023a). *Export area: European Union*. https://www.prosciuttodiparma.com/en/european-union/

Consorzio del Prosciutto di Parma (Parma). (2023b). *Preparation of Parma ham: Nutritional information*. https://www.prosciuttodiparma.com/en/preparation-of-parma-ham/

Cummings, R.G., and L.O. Taylor. 1999. Unbiased value estimates for environmental goods: A cheap talk design for the contingent valuation method. *American Economic Review*, 89, 649–65.

De Bauw, M., Franssens, S., & Vranken, L. (2022). Trading off environmental attributes in food consumption choices. *Food Policy*, 112, 102338. https://doi.org/10.1016/j.foodpol.2022.102338

De Bauw, M., Matthys, C., Poppe, V., Franssens, S., & Vranken, L. (2021). A combined Nutri-Score and 'Eco-Score' approach for more nutritious and more environmentally friendly food choices? Evidence from a consumer experiment in Belgium. *Food Quality and Preference*, 93, 104276. https://doi.org/10.1016/j.foodqual.2021.104276

Deselnicu, O. C., Costanigro, M., Souza-Monteiro, D. M., & Thilmany McFadden, D. (2013). A meta-analysis of geographical indication food valuation studies: What drives the premium for origin-based labels? *Journal of Agricultural and Resource Economics*, 38(2), 204–219. https://doi.org/10.22004/ag.econ.158285

Drugova, T., Curtis, K. R., & Akhundjanov, S. B. (2020). Are multiple labels on food products beneficial or simply ignored? *Canadian Journal of Agricultural Economics/Revue Canadienne d'agroeconomie*, 68(4), 411–427. https://doi.org/10.1111/cjag.12259

eAmbrosia. (2023). *The EU geographical indications register*. https://ec.europa.eu/info/food-farming-fisheries/food-safety-and-quality/certification/quality-labels/geographical-indications-register/

Egnell, M., Galan, P., Farpour-Lambert, N. J., Talati, Z., Pettigrew, S., Hercberg, S., & Julia, C. (2020). Compared to other front-of-pack nutrition labels, the Nutri-Score emerged as the most efficient to inform Swiss consumers on the nutritional quality of food products. *PLOS ONE*, 15(2), 1–18. https://doi.org/10.1371/journal.pone.0228179

Ernährungs Umschau (ErUm). (2013, February 14). *Mobilitätspanel: Frauen übernehmen Großteil der Haushalteinkäufe*. https://www.ernaehrungs-umschau.de/print-artikel/14-02-2013-mobilitaetspanel-frauen-uebernehmen-grossteil-der-haushaltseinkaeufe/

European Commission (EC). (2020). *Farm to Fork Strategy*. https://food.ec.europa.eu/horizontal-topics/farm-fork-strategy_en

European Food Agency (EFA). (2022). *Nutri-Score banned by Italian anti-trust authority*. https://www.efanews.eu/item/26073-nutri-score-banned-by-italian-anti-trust-authority.html

Eurostat. (2022). *Comparative price levels for food, beverages and tobacco.* https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Comparative_price_levels_for_food,_beverages_and_tobacco

Food and Agriculture Organization of the United Nations (FAO). (2021). *The nutrition and health potential of geographical indication foods*. Rome. https://doi.org/10.4060/cb3913en

Fortuna, G. (2022, October 14). Agrifood brief: Mediterranean (back)fire. *EURACTIV*. https://www.euractiv.com/section/agriculture-food/news/agrifood-brief-mediterranean-backfire/

Fortuna, G., Dahm, J., & Foote, N. (2022, February 11). Agrifood Brief: The last score that broke the label's back. *EURACTIV*. https://www.euractiv.com/section/agriculture-food/news/agrifood-brief-the-last-score-that-broke-the-labels-back/

Gassler, B., Faesel, C. K., & Moeser, A. (2022). Toward a differentiated understanding of the effect of Nutri-Score nutrition labeling on healthier food choices. *Agribusiness*, 39(1), 28–50. https://doi.org/10.1002/agr.21762

Gerini, F., Alfnes, F., & Schjøll. (2016). Organic- and animal welfare-labelled eggs: Competing for the same consumers? *Journal of Agricultural Economics*, 67(2), 471–490.

Gracia, A., & de-Magistris, T. (2016). Consumer preferences for food labeling: What ranks first? *Food Control*, 61, 39–46. https://doi.org/10.1016/j.foodcont.2015.09.023

Grunert, K. G., & Aachmann, K. (2016). Consumer reactions to the use of EU quality labels on food products: A review of the literature. *Food Control*, 59, 178–187. https://doi.org/10.1016/j.foodcont.2015.05.021

Gunlock, J. (2015, January 24). Food labeling fatigue. *Huffington Post*. https://www.huffpost.com/entry/food-labeling-fatigue_b_6215052

Hagmann, D., & Siegrist, M. (2020). Nutri-Score, multiple traffic light and incomplete nutrition labelling on food packages: Effects on consumers' accuracy in identifying healthier snack

options. *Food Quality and Preference*, 83, 103894. https://doi.org/10.1016/j.foodqual.2020.103894

Höhn, G. L., Huysmans, M., & Crombez, C. (2023a). Does terroir size matter? Protected geographical areas and prices of European hams. *Regional Studies*. https://doi.org/10.1080/00343404.2023.2187365

Höhn, G. L., Huysmans, M., & Crombez, C. (2023b). Healthy Food Traditions? Nutritional Quality and Food Composition of EU Geographical Indications. *Q Open*, 3(1), qoad014. https://doi.org/10.1093/qopen/qoad014

Huysmans, M., Noord, D. V., & Höhn, G. L. (2022). Do geographical indications certify origin and quality? A conceptual framework and an exploratory test on Gouda cheeses. *Worldwide Perspectives on Geographical Indications*. https://hal.science/hal-03791856

Julia, C., & Hercberg, S. (2017). Development of a new front-of-pack nutrition label in France: The five- colour Nutri-Score. *Public Health Panorama*, 3(4), 712–725.

Jumbo. (2022). *Gezonde producten*. https://www.jumborapportage.com/in-cijfers/gezondheid/gezonde-producten

Konrad, J. (2023, January 27). Discounter hängen Supermärkte ab. *Lebensmittel Zeitung*. https://www.lebensmittelzeitung.net/handel/nachrichten/consumer-index-der-gfk-discounter-haengen-supermaerkte-ab-169349?crefresh=1

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, 74(2), 132–157.

Lemken, D., Zühlsdorf, A., & Spiller, A. (2021). Improving Consumers' Understanding and Use of Carbon Footprint Labels on Food: Proposal for a Climate Score Label. *EuroChoices*, 20(2), 23–29. https://doi.org/10.1111/1746-692X.12321

Leufkens, D. (2018). The problem of heterogeneity between protected geographical indications: A meta-analysis. *British Food Journal*, 120(12), 2843–2856. https://doi.org/10.1108/BFJ-12-2017-0710

Lizin, S., Rousseau, S., Kessels, R., Meulders, M., Pepermans, G., Speelman, S., Vandebroek, M., Van Den Broeck, G., Van Loo, E. J., & Verbeke, W. (2022). The state of the art of discrete choice experiments in food research. *Food Quality and Preference*, 102, 104678. https://doi.org/10.1016/j.foodqual.2022.104678

Marnik, G. D. & van Heerde, H. J. (2023). Retailing in times of soaring inflation: What we know, what we don't know, and a research agenda. *Journal of Retailing*, (advance online publication). https://doi.org/10.1016/j.jretai.2023.07.002

McFadden, D. (1974). Conditional logit analysis of qualitative choice behavior. In P. Zarembka (Ed.), Frontiers in econometrics (pp. 105–142). Academic Press.

Menapace, L., & Raffaelli, R. (2020). Unraveling hypothetical bias in discrete choice experiments. *Journal of Economic Behavior and Organization*, 176, 416–430. https://doi.org/10.1016/j.jebo.2020.04.020

Moschini, G. C., Menapace, L., & Pick, D. (2008). Geographical indications and the competitive provision of quality in agricultural markets. *American Journal of Agricultural Economics*, 90(3), 794–812. https://doi.org/10.1111/j.1467-8276.2008.01142.x

Penn, J. M., & Hu, W. (2018). Understanding Hypothetical Bias: An Enhanced Meta-Analysis. American *Journal of Agricultural Economics*, 100(4), 1186–1206. https://doi.org/10.1093/ajae/aay021

Pérez y Pérez, L., Gracia, A., & Barreiro-Hurlé, J. (2020). Not Seeing the Forest for the Trees: The Impact of Multiple Labelling on Consumer Choices for Olive Oil. *Foods*, 9, 186. https://doi.org/10.3390/foods9020186

Qualivita. (2022). *AFIDOP, i formaggi DOP italiani uniti: Nutri-score ingannevole, penalizza prodotti certificati e sana alimentazione*. https://www.qualivita.it/news/afidop-i-formaggi-dop-italiani-uniti-nutri-score-ingannevole-penalizza-prodotti-certificati-e-sana-alimentazione/

Qualivita. (2023). *Rapporto Ismea-Qualivita 2022 sulle produzioni agroalimentari e vitivinicole italiane DOP IGP STG*. https://www.qualivita.it/osservatorio/rapporto-ismea-qualivita/#toggle-id-1

Rappoport, L., Peters, G. R., Downey, R., McCann, T., & Huff-Corzine, L. (1993). Gender and age differences in food cognition. *Appetite*, 20(1), 33–52. doi: 10.1006/appe.1993.1004

Reisch, L. A., & Zhao, M. (2017). Behavioural economics, consumer behaviour and consumer policy: State of the art. *Behavioural Public Policy*, 1(2), 190–206. https://doi.org/10.1017/bpp.2017.1

REWE. (2022). *Nutri-Score*. https://www.rewe.de/ernaehrung/nutri-score/

Retailtrends. (2013). *Consument niet prijsbewust, wel prijsgevoelig.* https://retailtrends.nl/news/32123/consument-niet-prijsbewust-wel-prijsgevoelig

Retailtrends. (2022). *Discounters winnen marktaandeel.* https://retailtrends.nl/news/67891/discounters-winnen-marktaandeel

Rijksoverheid. (2023). *Invoering Nutri-Score om gezonde voeding te stimuleren*. https://www.rijksoverheid.nl/onderwerpen/voeding/gezonde-voeding/nieuw-voedselkeuzelogo-nutri-score/invoering-nutri-score

Roquefort. (2022). *#Roquefortsansnutriscore*. https://www.roquefort.fr/nutri-score/

San Daniele. (2022). *San Daniele is healthy*. https://prosciuttosandaniele.it/en/the-ham/san-daniele-is-healthy/

Scheibehenne, B., Miesler, L., & Todd, P. M. (2007). Fast and frugal food choices: Uncovering individual decision heuristics. *Appetite*, 49(3), 578–589. doi: 10.1016/j.appet.2007.03.224

Scheibehenne, B., Greifeneder, R., & Todd, P. M. (2010). Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload. *Journal of Consumer Research*, 37(3), 409–425. https://doi.org/10.1086/651235

Sociaal en Cultureel Planbureau (SCP). 2019. *Zorg voor het huishouden en anderen*. https://digitaal.scp.nl/eenweekinkaart2/zorg-voor-het-huishouden-en-anderen/

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, 104746. https://doi.org/10.1016/j.foodqual.2022.104746

Statista. (2013). *Preissensibilität von Konsumenten in Deutschland und ausgewählten Ländern weltweit 2013*. https://de.statista.com/statistik/daten/studie/273141/umfrage/preissensibilitaet-von-konsumenten-in-deutschland-und-ausgewaehlten-laendern-weltweit/

Steenkamp, J. E. M. (1990). Conceptual model of the quality perception process. *Journal of Business Research*, 21, 309–333. https://doi.org/10.1016/0148-2963(90)90019-A

Stiletto, A., & Trestini, S. (2022). Is it really a piece of cake to label Geographical Indications with the Nutri-Score? Consumers' behaviour and policy implications. *PLOS ONE*, 17(11), e0277048. https://doi.org/10.1371/journal.pone.0277048

Török, Á., & Jambor, A. (2016). Determinants of the revealed comparative advantages: The case of the European ham trade. *Agricultural Economics – Czech*, 62(10), 471–481. https://doi.org/10.17221/177/2015-AGRICECON Determinants

Train, K. E. (2009). Discrete choice methods with simulation (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511805271

van Doorn, J., & Verhoef, P. C. (2011). Willingness to pay for organic products: Differences between virtue and vice foods. *International Journal of Research in Marketing*, 28(3), 167–180. https://doi.org/10.1016/j.ijresmar.2011.02.005

van Ittersum, K., Meulenberg, M. T. G., van Trijp, H. C. M., & Candel, M. J. J. M. (2007). Consumers' appreciation of regional certification labels: A pan-European study. *Journal of Agricultural Economics*, 58(1), 1–23. https://doi.org/10.1111/j.1477-9552.2007.00080.x

von Massow, M. (2019, August 12). We are reaching label fatigue. *Food Focus*. https://www.foodfocusguelph.ca/post/we-are-reaching-label-fatigue

Wardle, J., Haase, A. M., Steptoe, A., Nillapun, M., Jonwutiwes, K., & Bellisie, F. (2004). Gender differences in food choice: The contribution of health beliefs and dieting. *Annals of Behavioral Medicine*, 27(2), 107–116. https://doi.org/10.1207/s15324796abm2702_5

Appendices

Appendix 1: Descriptive statistics of sample and population

The sample (see Table A1.1) is not fully representative of the German or Dutch populations (see Table A1.2). It tends to overrepresent lower income groups, with about 60% of respondents reporting a below-average annual gross household income. In contrast, pensioners are underrepresented, with only 1% of respondents being retired. Although the proportion of participants with completed tertiary education is rather close to the general population, students are overrepresented at 33%. Younger age groups also see an overrepresentation, with 35% of respondents being under the age of 25. Nonetheless, our sample can still reveal useful insights because it covers respondents from major societal groups, is balanced regarding gender and all our respondents represent current consumers of ham.

Table A1.1 Descriptive statistics of sample

	Total (N=815)		**German (N=491)**		**Dutch (N=324)**	
Variable	**N**	**Share**	**N**	**Share**	**N**	**Share**
Age						
18-24 years	287	35%	163	33%	124	38%
25-49 years	485	60%	302	62%	183	56%
50-64 years	38	5%	24	5%	14	4%
> 65 years	5	1%	2	0%	3	1%
Gender						
Male	403	49%	239	49%	164	51%
Female	404	50%	247	50%	157	48%
N/A	8	1%	5	1%	3	1%
Education level (completed)						
Non-tertiary	544	67%	319	65%	225	69%
Tertiary	262	32%	167	34%	95	29%
N/A	9	1%	5	1%	4	1%
Household income (gross)						
< 50,000 p.a.	469	58%	293	60%	176	54%
≥ 50,000 p.a.	248	30%	150	31%	98	30%
N/A	98	12%	48	10%	50	15%
Employment						
Employed	475	58%	292	59%	183	56%
Unemployed	37	5%	20	4%	17	5%
Student	273	33%	161	33%	112	35%
Retired	7	1%	4	1%	3	1%
N/A	23	3%	14	3%	9	3%

Table A1.2 Descriptive statistics of German and Dutch population (2020)

Variables	***Germany*** *Population of 83,166,711*	***The Netherlands*** *Population of 17,407,585*
Age		
15-24 years	10%	12%
25-49 years	31%	32%
50-64 years	23%	21%
> 65 years	22%	20%
Education level (completed)		
Non-tertiary	73%	63%
Tertiary	27%	37%
Household gross income (average)	56,580 €	75,200€
Employment		
Employment rate (share of ages 15-64)	74%	79%
Students (share of total population)	4%	5%
Retired (share of total population)	28%	25%

Source: Own elaboration based on Eurostat data.

Appendix 2: Preference heterogeneity among consumers

There is a clear indication of preference for the PDO and better Nutri-Score on average. However, the significant standard deviations of all our random coefficients in Models 1 and 2 reveal that there is indeed heterogeneity regarding preferences in our sample, i.e. some respondents value the characteristic more than others. Table A2 reports the standard deviations of the random coefficients.

Table A2 Standard deviations of random coefficients

	Model (1)	Model (2)
No-purchase (ASC)	3.28***	3.28***
	(0.16)	(0.17)
PDO Parma ham (GI)	2.77***	2.76***
(base level: generic)	(0.12)	(0.13)
Nutri-Score D (NS)	2.52***	2.51***
(base level: E)	(0.11)	(0.12)
*GI*NS*		*0.99***
(interaction)		*(0.29)*
Summary Statistics		
N	22,005	22,005

Notes: *p < 0.1, **p < 0.05 and ***p < 0.01. Standard errors are shown in parentheses.

Given the observed preference heterogeneity in our sample we control for socio-demographic variables by interacting them with the *GI* and *NS* variables in Models 3, 4, 5 and 6 (see Table A3). First, there may be a difference between male and female respondents as some research highlighted gender differences regarding diet purchases. For example, some studies find that women are more likely to avoid unhealthy foods (Wardle et al., 2004). Traditionally, women also tend to be responsible for household tasks such as grocery shopping in Germany and the Netherlands (ErUm, 2013; SCP, 2019). Thus, while our sample is representative in terms of gender, women's preferences may matter more regarding the actual purchases and premiums. Hence, we introduce interactions in Model 3 with the dummy *Male* that takes the value of 1 if the respondent is male. On the one hand, we do not find that males are significantly less likely to value a better Nutri-Score compared to females. On the other hand, males seem to be more prone to label fatigue and are more likely to value PDO Parma ham. Thus, our estimations may overestimate the premium for Parma ham as women tend to indicate lower WTP.

In contrast, our analysis may underestimate the premium for PDO Parma ham or a Nutri-Score due to the overrepresentation of respondents in the youngest age group. Hence, we use the variable *Young*, which takes the value of 1 if the respondent is 24 or younger, for the interactions in Model 4. While there is a slight indication for lower WTP for PDO Parma ham, there is no statistical difference concerning WTP for the Nutri-Score D or the interaction effect. Nonetheless, the premium for PDO Parma ham may be underestimated given the overrepresentation of younger respondents and students.

Next, as we have a bi-national sample, we control for potential differences between German and Dutch consumers by interacting in Model 5 with the dummy variable *Dutch*, which takes the value 1 if the respondent is Dutch. German and Dutch consumers tend to be price sensitive regarding foods, which may have been amplified by higher inflation (Marnik et al., 2023). Additionally, despite their economic strength, Germany and the Netherlands have relatively low consumer price levels (Eurostat, 2022). Thus, our estimations of marginal WTP probably represent an underestimation compared to other EU countries. In our sample, Dutch respondents have a lower WTP for PDO Parma ham and the better Nutri-Score D compared to German respondents. In addition, the Dutch respondents seem more prone to label fatigue, i.e. the combined effect of two positive labels is less than the sum of the labels separately, because the interaction of *Dutch* and *GI***NS* is negative and significant, but again at the 10% level only.

Finally, Stiletto and Trestini (2022) found that respondents with prior knowledge about the PDO label can differ in their preferences. In our sample, the awareness of the PDO label prior to our experiment was with 25% indeed low. Therefore, we include interactions in Model 6 with the variable '*GI-prior*' which takes the value of 1 if the respondent knew about the PDO label beforehand. The interaction of *GI-prior* and *GI* is positive while the interaction of *GI-prior* and *NS* is negative. As expected, consumers with prior knowledge of the PDO label have a stronger preference for Parma ham and a weaker preference for the Nutri-Score D. The three-way interaction of *GI-prior* with *GI* and *NS* is again negative and insignificant.

All in all, the preference heterogeneity in our sample is partially explained by gender, age, nationality and prior knowledge of the PDO label. Overall, the sample of rather young German and Dutch consumers that tend to be unaware of the PDO label is likely to underestimate the WTP for PDO Parma ham and a better Nutri-Score D. Nonetheless, our results are still insightful as the sample represents current consumers of ham in two major export markets of PDO Parma ham where the Nutri-Score was adopted or about to be adopted.

Table A3 Mixed logit regressions controlling for socio-demographics

	Model (3)	Model (4)	Model (5)	Model (6)
Price	-3.18***	-3.15***	-3.16***	-3.16***
	(0.09)	(0.09)	(0.09)	(0.09)
No-purchase (ASC)	-12.48***	-12.35***	-12.36***	-12.39***
	(0.35)	(0.35)	(0.35)	(0.35)
PDO Parma ham (GI)	1.96***	2.43***	2.48***	2.12***
(base level: generic)	(0.18)	(0.17)	(0.18)	(0.15)
Nutri-Score D (NS)	1.72***	1.47***	1.66***	1.68***
(base level: E)	(0.18)	(0.16)	(0.17)	(0.15)
*GI*NS*	0.04	-0.31*	-0.01	-0.29*
(interaction)	(0.18)	(0.18)	(0.17)	(0.16)
*GI*Male*	0.60**			
	(0.24)			
*NS*Male*	-0.44*			
	(0.23)			
*GI*NS*Male*	-0.47*			
	(0.26)			
*GI*Young*		-0.45*		
		(0.26)		
*NS*Young*		0.10		
		(0.29)		
*GI*NS*Young*		0.25		
		(0.27)		
*GI*Dutch*			-0.52**	
			(0.25)	
*NS*Dutch*			-0.40*	
			(0.24)	
*GI*NS*Dutch*			-0.52*	
			(0.27)	
*GI*Prior*				0.68**
				(0.30)
*NS*Prior*				-0.63**
				(0.27)
*GI*NS*Prior*				0.21
				(0.31)
Summary Statistics				
N	21,789	22,005	22,005	22,005
Log-likelihood	-5010.91	-5075.85	-5068.89	-5066.57
AIC	10051.83	10181.7	10167.77	10163.14

Notes: *$p < 0.1$, **$p < 0.05$ and ***$p < 0.01$. Standard errors are shown in parentheses.