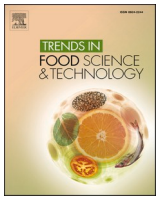
Trends in Food Science & Technology 170 (2026) 105435



Contents lists available at ScienceDirect

Trends in Food Science & Technology

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



To label or not to label? A meta-regression analysis of consumer willingness to pay for sustainability claims

Giulia Andreani [a], Juan Tur-Cardona [b,*], Federico Antonioli [c], Jesus Barreiro-Hurle [c], Giovanni Sogari [a], Davide Menozzi [a]

[a] Department of Food and Drug, University of Parma, Parma, Italy
[b] European Commission, Joint Research Centre (JRC), Ispra, Italy
[c] European Commission, Joint Research Centre (JRC), Seville, Spain

ARTICLE INFO

Handling Editor: Dr. S Charlebois

Keywords:
Meta-regression analysis
Publication bias
Sustainable consumption
Sustainability label
Willingness-to-pay

ABSTRACT

Background: Sustainability labels (SLs) are increasingly promoted to guide consumer behavior towards more sustainable consumption choices. As a result, research on SLs has increased with numerous studies and reviews highlighting how specific sustainability claims are valued by consumers.

Scope and approach: This study presents a comprehensive meta-analysis of willingness to pay (WTP) for SLs displayed on food products, including a wide range of studies focusing on diverse food categories. To conduct the meta-analysis on SLs, Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines was used.

Key findings and conclusion: The analysis reveals that consumers, on average, are willing to pay a 29 % price premium for products bearing SLs. This mean value masks significant heterogeneity in relation to label and product categories as well as consumer socioeconomic factors. Notably, carbon labels were associated with lower percentage premium WTP, while geographic and animal welfare labels drove higher premiums. From a product perspective, higher percentage premium WTP is observed for SL when used in meat, dairy, alcohol, coffee, and cereals when compared to vegetables. Moreover, consumers in higher-education segments tend to exhibit lower relative WTP, potentially reflecting greater skepticism toward label claims. Percentage WTP premium also decreases with higher product prices, indicating price sensitivity. Studies published after 2017 tend to report higher percentage WTP premium values, possibly linked to increased awareness of sustainability issues. The results empower policymakers, marketers, and stakeholders seeking to design more effective sustainability communication strategies, emphasizing the importance of label type, context, and consumer segmentation.

1. Introduction

Sustainability has become a critical priority for national governments and international organizations. In 2015, the 2030 Agenda for Sustainable Development was launched by the United Nations as a global call to action to achieve 17 Sustainable Development Goals (SDGs), aiming to foster environmental, social, and economic sustainability at a global scale (European Commission, 2020; UN, 2015). To achieve these goals, the transformation of food systems has played a central role since the beginning, including the need that diets are not only inclusive, healthy, and nutritious but also environmentally sustainable. Given the increasing global population, shifts in consumption patterns, climate change, and the depletion of natural resources, all stakeholders in the food system must be equipped with the necessary tools to contribute to and drive meaningful changes (European Commission, 2020; FAO F. and A. O. of the U. N, 2018).

When targeting consumers, various strategies can be implemented to make individuals make more sustainable purchasing choices, ranging from subtle interventions, such as nudging (Tirion et al., 2025), to more direct and intrusive measures, as market-based interventions and legislative changes (Ammann et al., 2023). While the latter have proven to be effective, they also often face resistance due to lower consumer acceptance (Ammann et al., 2023) and lack of political will (see Bødker et al., 2015). Among the less intrusive approaches, the use of sustainable

* Corresponding author.
E-mail address: juan.tur-cardona@ec.europa.eu (J. Tur-Cardona).

https://doi.org/10.1016/j.tifs.2025.105435
Received 22 August 2025; Received in revised form 6 November 2025; Accepted 9 November 2025
Available online 11 November 2025
0924-2244/© 2025 The Authors. Published by Elsevier Ltd. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

labels (SLs)[1] on food packaging has emerged as a promising strategy to guide consumers toward more sustainable decisions (Li & Kallas, 2021; Potter et al., 2021).

Over recent decades, various SLs have been developed by both private actors of the food system, non-governmental organizations and public institutions, and are currently present on food products available in retail stores, often focusing on a single specific sustainability aspect (Dendler, 2014). As a result, research on sustainability labels has increased significantly, with numerous studies and reviews highlighting whether and how such specific sustainability aspects (e.g., carbon footprint, animal welfare or use of pesticides, among others) are valued by consumers and how they can potentially influence their behavior on the final consumer market (Abdu & Mutuku, 2021; Rondoni & Grasso, 2021; Thøgersen, 2023). Results generally show that consumers show positive attitudes toward SLs and, overall, their willingness to pay (WTP) for food products bearing sustainability-related claims increases (Rondoni & Grasso, 2021). However, the outcome can vary significantly based on consumer sociodemographic characteristics and on the type of label, particularly when consumers face trade-offs between sustainability and other product characteristics (such as quality or origin), situations that weaken the importance of SLs (Rondoni & Grasso, 2021; Thøgersen, 2023). To address this issue, the idea of including multiple sustainability dimensions in a single and comprehensive meta-sustainability label has been proposed (Dendler, 2014; Sonntag et al., 2023; Torma & Thøgersen, 2021). Such a label would reduce the noise of multiple individual labels while providing consumers with a clearer, more holistic understanding of a product's sustainability. Despite this promising solution, as highlighted in prior research (Dendler, 2014; Sonntag et al., 2023; Torma & Thøgersen, 2021), meta-sustainability labels remain underexplored, with further investigation necessary to support their effective implementation.

The development of SLs and adherence by producers to their standards imply that food products bearing a SL have to incur additional costs (e.g., financial investments for new equipment, change in inputs, transaction costs for switching to different production technologies, and administrative costs of certification). These additional costs eventually have to be passed through to increase the final consumer market (Yenipazarli, 2015). Therefore, understanding the extent of consumers' WTP is essential to assess the economic viability of any labelling scheme: a significant market demand for more products with a sustainability label would encourage companies to invest in undertaking sustainability efforts to reach the SLs requirements. On the contrary, the absence of consumers' WTP for a premium would suggest the application of alternative approaches to labelling for promoting sustainability in the food system.

The strength of our work lies in its comprehensive, extensive, up-to-date meta-analysis of SLs, which goes beyond previous reviews that often focused on a single label type or food category. Indeed, by encompassing a broad spectrum of SLs across various products and extending the analysis beyond discrete choice experiments (DCEs), our study provides a more nuanced perspective on consumer choices and intentions, capturing diverse purchasing contexts and interactions with other product attributes. This approach allows us to identify key factors influencing consumer responses, assess the economic feasibility of sustainability labelling, and highlight conditions under which SLs are most effective. Ultimately, our findings offer new actionable insights into the effectiveness of SLs in driving more sustainable consumer choices and purchasing intentions, also providing industry stakeholders with a more impactful and market-aligned sustainability labelling strategy.

2. Comparison with other meta-analyses

Previous reviews summarized and compared studies on different SLs, concluding on their generally positive behavioral outcomes by steering dietary choices toward greener alternatives (Majer et al., 2022; Potter et al., 2021). Some of the most comprehensive work was conducted through meta-regression analysis, offering an empirical assessment of the impact of SLs across a variety of labels, food products, and consumer segments (Table 1). Nevertheless, these studies differ significantly in terms of scope, conceptual framework, methodology, and the level of granularity of the explored heterogeneity. Several of these reviews are focused on consumers' WTP a premium for a single dimension of sustainability, such as animal welfare claims (Clark et al., 2017; Lagerkvist & Hess, 2011), geographical origin labels (Santeramo & Lamonaca, 2020; Yeh & Hirsch, 2023), eco- or environmental-friendly labels (Abdu & Mutuku, 2021; Bastounis et al., 2021) or health and nutrition labels (Dolgopolova & Teuber, 2018; Song et al., 2021), providing domain-specific insights. It is evident that, in those cases, our study integrates and refines the analysis for specific sustainability features, such as animal welfare, geographical origin label and eco- or environmentally friendly labels with other sustainability dimensions into a more consistent and larger meta-analysis offering a wider cross-attribute analysis.

Other studies, such as Yang and Renwick (2019) and Abdu and Mutuku (2021), provided a meta-analysis on various sustainability labels but were limited to a specific product category, such as, respectively, animal and coffee products. In these cases, our analysis broadens the scope to encompass all food categories, integrating both livestock and plant-based products within a unified framework that examines the effects of diverse sustainability labels. This approach can provide actionable insights for farmers and policymakers to better tailor sustainability communication across different supply chains.

Finally, up to the authors' knowledge, very few studies provide multiple-label and product categories analysis related to sustainability, namely Tully and Winer (2014), Li and Kallas (2021) and Piracci et al. (2024).

Tully and Winer (2014) analyzed the effect of different socially responsible labels benefiting animals (e.g. animal welfare labels), humans (e.g., fair wages or good working conditions), or the environment (e.g. protect rainforests, reduce pollution, etc.), encompassing a wide range of products (e.g. wood, clothing, etc.), therefore preventing a more detailed analysis on specific food categories.

Li and Kallas (2021) conducted one of the first large-scale quantitative meta-analyses of consumers' WTP for sustainable food labels; covering 80 studies, they report an average WTP premium of 29.5 %. On the sustainability attributes, the authors conducted the analysis on the aggregate (e.g., carbon label and other eco- and sustainability labels), not considering nutrition and geographical origin labels. In this regard, our analysis updates, expands and refines the WTP estimates, offering a disaggregation by label type, including nutrition and geographical labels, and offering a wider regression on multiple moderators, adopting a strict PRISMA protocol and relying on a richer dataset of observations.

Piracci et al. (2024) provided the most recent and conceptually sophisticated meta-analysis of sustainability labelling. Their study analyses environmental and social labels, highlighting the heterogeneity of price premiums across label types, and showing weaker consumer valuation of social labels relative to environmental ones. Their work, although similar to ours, focuses on the effects of broader label categories divided according to three different criteria (i.e., broadness of formulation, sustainability dimension or scope and information provided), while the results we obtained allow for distinguishing the effect of labels specific for a given trait of the sustainability dimension (e.g., carbon label). Moreover, the study focusses specifically on the distribution of the WTP among different sustainability-label categories and, despite including moderator effects in their model specification (such as product categories, methodological issues, etc.), these are not explicitly discussed in terms influence on their influence on the WTP of the sustainability labels. Therefore, our study advances beyond Piracci et al. (2024) by providing more recent and disaggregated results, including in

[1] The terms label and claim are used interchangeably throughout the study.

Table 1
Overview of previous meta-analyses.

Study	Sustainability dimensions	Label(s) and scope	Methodological approach	Time coverage	Overlap of studies	Key results	Identified gaps and limitations
Lagerkvist and Hess (2011)	Animal welfare	Effect of various farm animal welfare attributes (e.g. Indoor light change, rest areas, etc.).	Systematic review and meta-analysis of 24 studies; 106 WTP estimates; meta-regression on socio-demographics.	1995–2008	0/24	Consumers have positive WTP for farm animal welfare; WTP varies by income, age (negative), and information.	The study is limited to animal welfare dimension, little or no focus on sustainability and eco- aspects.
Tully and Winer (2014)	Animal welfare, Eco-label Fair Trade & labour label	Socially responsible products (environmental, human, animal beneficiaries).	Literature review and meta-analysis on 80 studies; 163 WTP estimates; meta-regression models.	1996–2012	3/80	Average premium 16.8 %; higher when benefits accrue to humans.	Aggregates heterogeneous labels; not specific to food products, no cross-category analysis; not systematic review applied.
Clark et al. (2017)	Animal welfare	Effect of farm animal welfare on WTP; specific focus on reduced incidence of animal production diseases.	Systematic review and meta-analysis of 54 studies; 335 WTP estimates; random-effects meta-analysis.	1996–2014	1/54	Small but positive WTP for improved farm animal welfare; heterogeneity by animal type and region.	The study is focused on reduced animal diseases; limited attention on sustainability context; no % premiums.
Dolgopolova and Teuber (2018)	Health & Nutrition label	Health claims, focus on "lowering cholesterol" claim	Literature review and meta-analysis of 36 studies; 186 WTP estimates; meta-regression models.	2003–2016	0/36	The health claim of "lowering cholesterol" has a significantly positive influence on WTP	Focus on health claims, not sustainability; not systematic review applied.
Yang and Renwick (2019)	Animal welfare, Geographical label, Organic, Eco-label and Other labels	Different labels in livestock products: animal welfare, geographical origin, organic, eco-friendly, and safety.	Systematic review and meta-analysis of 94 studies; 555 WTP estimates; meta-regression models.	1998–2017	1/94	WTP varies across products (beef/dairy > lamb), attribute (from 20 % traceability to 36 % organic), and region; increasing WTP over time.	Study limited to livestock products; not extended to plant-based or other general foods.
Santeramo and Lamonaca (2020)	Geographical label	Geographical origin and geographical indication labelling.	Systematic review and meta-analysis of 27 studies; 194 WTP estimates; meta-regression on method, product and country.	2001–2018	0/27	Relevance of geographical labels varies by product type and country.	Focused on geographical origin label, no link with sustainability framework.
Li and Kallas (2021)	Eco-label, Geographical label, Organic, Fair Trade & labor label, Sustainability label	Sustainable food attributes (environmental-friendly, fair trade, organic and local) worldwide and their heterogeneity.	Literature review and meta-analysis on 80 studies; subgroup and meta-regression analyses.	2000–2020	7/80	Average WTP 29.5 %; influenced by gender, region, method, and food category; WTP value of organic attribute higher than other attributes.	Aggregates across sustainability attributes (e.g. carbon label); nutrition and geographical origin not considered; no strict protocol applied for systematic review.
Bastounis et al. (2021)	Eco-label, Organic, Sustainability label	Ecolabels for environmentally sustainable foods; effects of intervention characteristics and demographic variables.	Systematic review and meta-analysis of 35 studies; 129 WTP estimates; meta-regression models.	2002–2019	6/44	WTP premium higher for organic labels; women and less educated show higher WTP.	WTP expressed in PPP $/kg limits comparability; the analysis focuses on aggregated environmental labels only.
Abdu and Mutuku (2021)	Eco-label, Fair Trade & labor label, Geographical label, Organic	Coffee labelling analysis: organic, fairtrade, eco-label, country-of-origin.	Systematic review and meta-analysis on 22 studies; 97 WTP estimates; meta-regression models	2005–2019	3/22	Positive WTP for eco-coffee, especially organic; method type affects estimates.	Product-specific study (coffee); limited generalizability; lacks health and welfare dimensions.
Song et al. (2021)	Health & Nutrition label	Nutrition and different type of labels (interpretive, warning and color-coded).	Systematic review and network meta-analysis on 134 studies; 156 WTP estimates; meta-regression models.	1990–2021	3/120	Both traffic light and warning labels improve healthy choices; lab experiments dominate.	Focus on nutrition, not sustainability; limited external validity.
Yeh and Hirsch (2023)	Geographical label	Country-of-origin labelling (COOL).	Systematic review and meta-analysis of 59 studies; 204 WTP estimates; meta-regression models.	2007–2021	8/59	Positive WTP for COOL (average 26 % premium); with regional and studies' choice design significant variability.	Single-attribute focus; not connected with other sustainability goals.
Piracci et al. (2024)	Animal welfare Carbon label Eco-label Fair Trade & labor label	The original labels were aggregated into categories according to three different criteria: broadness of formulation (specific or generic),	Systematic review and meta-analysis of 110 studies; 648 WTP estimates; meta-regression models.	2006–2021	15/110	Premiums for SLs are highly heterogeneous; social issues (fairness) less valued; both generic and specific claims accepted, heterogeneity	Focuses on broader labels, does not allow to distinguish the effect of specific ones (e.g. carbon label); moderating effects of other variables not

(*continued on next page*)

Table 1 (*continued*)

Study	Sustainability dimensions	Label(s) and scope	Methodological approach	Time coverage	Overlap of studies	Key results	Identified gaps and limitations
	Organic Sustainability label	sustainability dimension (environmental, social or combination) or scope and information provided (environmental, production method, animal welfare, social fairness and sustainable production labels).				of results explored by product price range (low-priced vs. high price).	discussed (products, time, methods, etc.).

the analytical framework socio-economic, product-categories, temporal and methodological moderators. Overall, under the methodological angle, we apply the publication-bias-robust PET-PEESE estimator, which has been widely proven to provide superior estimates to the conventional fixed and random effects MRA, even when there are no publication bias effects (Stanley, 2017; Stanley & Doucouliagos, 2014b, 2015).

Although most meta-analyses address the sustainability of food products, the studies are based on different databases with little overlap (below 15 % for most included meta-analyses) due to their different scopes, specific objectives, and review protocols. This leads to complementary findings in the different studies covering the sustainability of food products with a different approach.

In general, most of the meta-analysis studies considered concluded that SLs are a viable strategy for retailers, suggesting that these should be considered in an integrated policy approach to promote more sustainable food choices. However, the need for updated meta-analyses remains strong, especially as the number of publications on the topic continues to grow in response to global sustainability concerns (Piracci et al., 2024). Moreover, in all the reported meta-analyses, the time coverage, i.e. the year of publication of the original studies, is prior with respect to our study, which therefore provides a more accurate and up-to-date picture. Therefore, our work contributes to this topic by providing a refined meta-analysis on WTP for sustainability attributes, with an extensive multi-language (i.e., English, Italian, French and Spanish) coverage of 173 studies and more than 1000 observations, covering food products and nine label categories.

3. Methodology

To conduct the meta-analysis on SLs, we followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines. The protocol of the present research was previously registered on OSF (https://doi.org/10.17605/OSF.IO/E9MKR), which includes the publication of the PRISMA checklist addressing all stages of our work, from the objectives and the introduction definition to the analysis. By following open science best practices, we contribute to the extension of best practices in economic research (Arpinon & Lefebvre, 2025; Christensen & Miguel, 2018; Finger et al., 2025) which is still far from common practice among social scientists (Ferguson et al., 2023).

3.1. Search strategy

The primary intervention considered in our research is sustainable food labeling, which includes all types of labels providing information on one or more sustainable characteristics of a food product, ranging from those existing and widespread labels, such as organic production or fair trade, more niche ones and even hypothetical ones (e.g., low water usage or free from pesticides) (Table A4). The outcome of interest was the effect on consumer willingness to pay (WTP). A comprehensive search strategy was developed by the researchers involved in the study with keywords that covered five different dimensions: 1) the specific mention of sustainability, 2) the inclusion of at least one sustainability aspect (e.g., social, environmental, economic, health, etc.), 3) the reference to food or specific food categories, 4) referred to consumers as study subject, and 5) the included valuation tool (e.g. label, claim, certificate, etc.). A detailed description of the terms used to cover these concepts is reported in the Appendix (Table A1). From these, the final keyword search strategy was defined (Appendix, Table A2) and, on March 20, 2023, two electronic databases (Scopus and Web Of Science) were systematically searched. Identified articles were then filtered to consider only original articles published after 1990 and written either in English, Italian, Spanish, or French. Additional primary exclusion criteria were also used to refine our search – all filters used in the search strategy are reported in the Appendix (Table A3). A total of 5150 articles were retrieved, and 423 duplicates were removed using Covidence as software (Covidence systematic review, 2020).

3.2. Eligibility criteria

The identification of eligible studies was performed using a 3-step approach: i) a title and abstract screening for which eligibility was assessed independently by two investigators and discrepancies were resolved by a third investigator, ii) a full-text screening of the remaining studies, carried out independently by 6 researchers, and iii) data extraction, conducted independently by 6 researchers.

The sample included original research papers addressing sustainable food purchase intentions or behavior, that included as well consumer valuations for sustainability attributes signaled with a label. As pre-registered, studies were excluded considering the following criteria:

- Studies were excluded when their focus was not on sustainable food purchase intention or behavior. For instance, studies excluded were those assessing the sustainability impact of different types of diets (i.e. vegan or vegetarian diet) or studies on technological assessment (e.g. studies detailing the sustainable impacts of agricultural or production practices of products) unless linked with specific sustainable food products consumer valuation or purchase. Similarly, research on food behaviors misaligned with the study's objectives (e.g., weight loss outcome, pills, dietary supplements, etc.), was also excluded to maintain relevance to the core topic of sustainable food choices.
- Studies primarily involving consumers as subjects were selected, excluding those focusing on other stakeholders like farmers.
- Studies with no reference to sustainability label elements (e.g., sustainable standard, logo, label, information, etc.) were excluded. Also those studies without assessments of the cause-effect relationship between the sustainability label element and the variable of interest (WTP) were excluded.
- Full text not in English, French, Spanish, or Italian: studies in other languages were excluded as these are the languages the authors can understand confidently.
- Only original research articles were included, excluding reviews or conference proceedings. When the full paper was not retrieved, even after contacting the corresponding author, the study was excluded.

After the title and abstract screening and the full-text screening, 173 papers were considered for the data extraction. Comprehensive information on the screening process is shown in the PRISMA flowchart in Fig. 1.

For data extraction, we used a pre-tested and pre-approved form developed by the authors to ensure consistency. The form included information on the food categories of the products under investigation, the types of labels, the socioeconomic characteristics of the sample, and additional variables related to methodological and publication details. Further details on the variables and their levels are provided in the Independent Variables in the meta-regression analysis (MRA) section.

3.3. Meta-regression analysis

The present study aims to deliver an unbiased and comprehensive empirical analysis assessing the effect of SLs of food products on consumers' WTP to provide evidence-based insights on the overall potential impact of sustainable labelling on the final consumer market. Given the vast number of empirical applications, MRA represents an appropriate tool for modelling, estimating and explaining the heterogeneity of reported WTP results. In the words of Stanley (2008, p. 105) "[…] *MRA can help researchers see through the murk of random errors and selected misspecification biases to identify and magnify the underlying statistical structure that characterizes genuine empirical effects*". MRA entails a regression analysis of existing empirical results on the drivers of heterogeneity (e.g., geographical scope of the study, the year, sample characteristics, the sustainability labels assessed, the attributes considered, and the econometric method applied, among others), as well as the scope and precision (i.e., the standard errors) of the estimated WTP. This allows to test the sensitivity of the parameter of interest to given objective characteristics, controlling for systematic heterogeneity and publication bias and obtaining an 'estimate of estimates', hence delivering a unique and objective overall assessment (Costa-Font et al., 2011; Stanley, 2001, 2005, 2008; Stanley & Doucouliagos, 2012).

Publication bias is possibly the main challenge when it comes to MRA: empirical results that are statistically significant and larger in magnitude are often easier to publish, generating the over-representation of larger and more significant effects (Stanley, 2008). Such emphasis on statistical significance may lead to p-hacking or data dredging, manipulating the hypothesis, as well as the model specification, fishing for significant results (e.g. multiple hypothesis testing without adjusting statistical inferences) and/or those confirming the theory by having the right sign, incentivizing selective reporting (Ferraro & Shukla, 2023). As a result, manuscripts with null or statistically non-significant results tend to remain in the 'file drawer', causing publication selection and biasing the average reported empirical effect of the literature away from zero (Stanley, 2008, 2017; Stanley & Doucouliagos, 2015). Given the recent attention scholars have paid to data dredging and p-hacking in agricultural economics (Ferraro & Shukla, 2023; Josephson & Michler, 2018; Rommel & Weltin, 2021), we particularly account for such potential publication bias in our econometric model.

3.3.1. Dependent variable in the MRA

The focus of our work is on the estimates of consumers' WTP for food products bearing a sustainable label. Nevertheless, as a result of the wide geographical and product-typology scopes of the meta-analysis, the WTP estimates are expressed in different currencies, different years, units of measure, and for different products, entailing significantly different values (e.g., wine versus tomatoes) (Oczkowski & Doucouliagos, 2015), and need to be comparable (Havránek et al., 2020; Oczkowski & Doucouliagos, 2015). As per common practice (Deselnicu et al., 2013; Lagerkvist & Hess, 2011; Lusk et al., 2005; Piracci et al., 2024; Printezis et al., 2019; Santeramo & Lamonaca, 2021; Yeh & Hirsch, 2023), we define our variable of interest using the relative percentage premium as a dependent variable, as the relative surcharge in percentage terms over some reference price[2]:

$$\%WTP = \frac{\left(WTP_{sustainable\ product} - P_{counterpart}\right)}{P_{counterpart}}$$

3.3.2. Independent variables in the MRA

Independent variables should capture the underlying heterogeneity of our variable of interest (Havránek et al., 2020). Accordingly, we use the information on the nature of the label, the food main categories to which the product belongs, the socioeconomic characteristics of the targeted population and additional variables capturing methodological and publication characteristics. The different labels found were classified in homogeneous categories according to their basic scope and the informative content (Piracci et al., 2024). Particularly, labels were grouped into the following categories: sustainability labels (for labels including an overarching concept of sustainability), animal welfare, eco-labels (including labels relating to water usage, waste management, pesticide residues soil or generic environmental labels), fair labels (including Fairtrade, social responsibility, fair labor or direct trade), geographical origin, health or nutrition disclaims, organic and other label categories (such as traceability, freshness, quality, farming practices or GM content) (see Table A4). In the same way, homogeneous food categories were defined: vegetables, meat products, milk, honey, eggs, fish products, coffee and tea, cereals, fruits and other foods (e.g., ready-made meals, generic food products or restaurant meals). Categories were defined in such a way that groups were homogeneous and balanced in terms of observation size, avoiding categories populated by very few observations, which may yield non-generalizable insights. The reference price[3] of the products in the studies was also included as an independent variable, expecting that respondents might be more sensitive to price increases on products with higher prices (Yang & Renwick, 2019). Despite referring to different sizes and quantities of products, such a variable is expected to anchor respondents' perception, which may be relevant when studying the effect of the relative percentage premium price.

Independent variables also included socioeconomic information about the sample of the population and the characteristics of the country of study. These were average values for age, percentage of females and university education of the sampled populations. The share of food expenditure in the country of the study was also used as a proxy for the importance of food in economic terms, accounting for the income effect of the region of study. Similarly, in other studies, measures macroeconomic variables such as GDP (Yang & Renwick, 2019) or geographic region (Abdu & Mutuku, 2021; Li & Kallas, 2021; Yeh & Hirsch, 2023) are used to indicate the potential regional differences in the study context. However, in our study, the variables with geographic region, GDP and share of food expenditure are correlated; thus, we prioritize food share expenditure.

Other variables included information on the methodological aspects of the studies. Including different labels in the same study might also affect WTP estimates, either reducing the WTP for the individual labels due to the overcrowding effect on consumers, or it might have positive effects by triggering perceptions about the food characteristics. Similarly, the inclusion of attribute levels can also affect the results (e.g., Caussade et al., 2005; Ryan & Wordsworth, 2000). Therefore, we controlled for the number of labels and the number of levels of the sustainable labels included in each study. Likewise, evidence suggests

[2] When the reference price was not reported the average value of all pricing points defined in the price attribute was used.

[3] All prices were converted to euro. When the study used a different currency, prices were converted to euro using purchasing power parity (PPP), which is a conversion rate that reflects the relative cost of goods and services in different countries, essentially indicating how many units of one currency are needed to buy the same amount of goods as one unit of another currency.

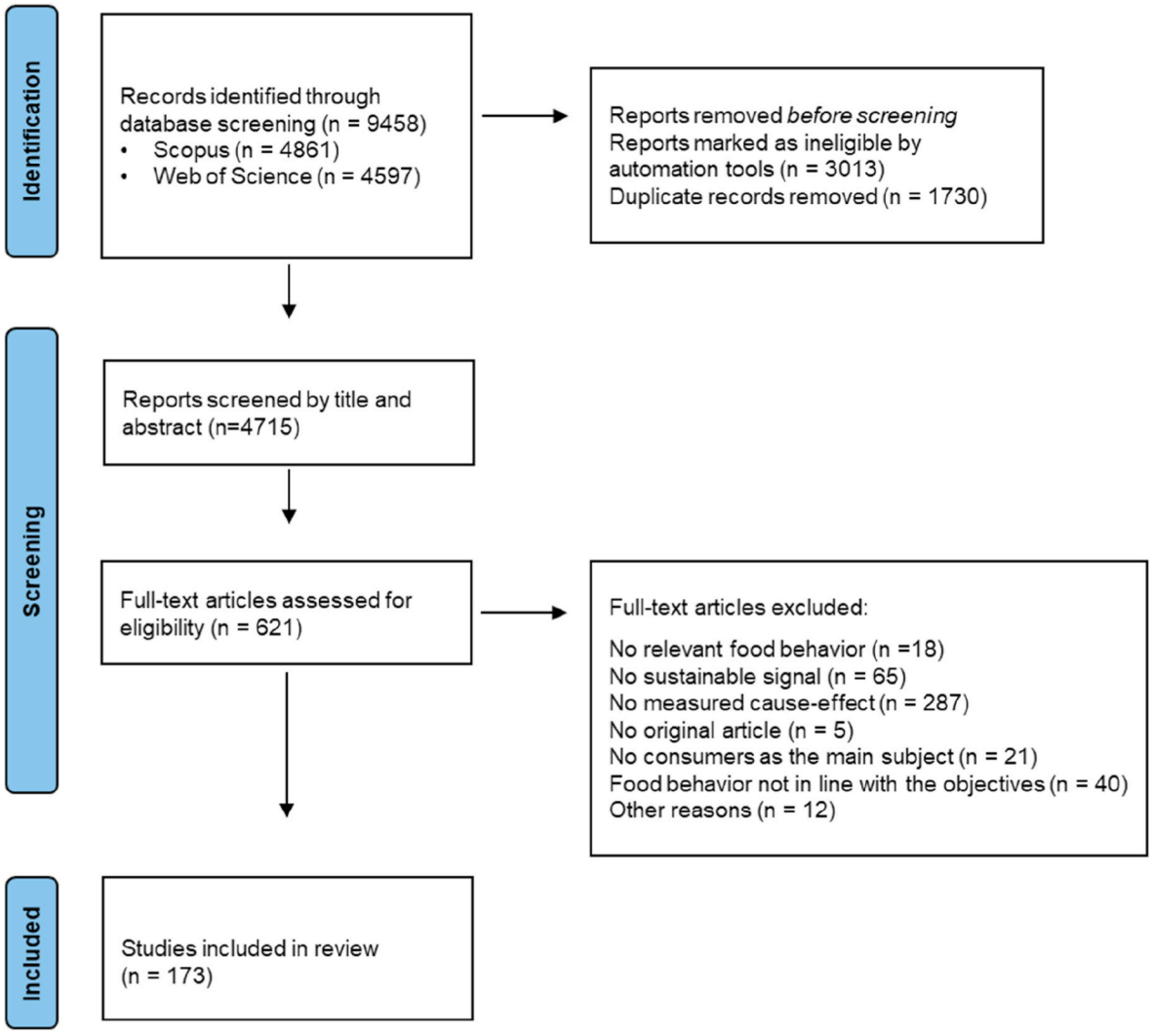


Fig. 1. PRISMA flowchart.

that differences might arise between online surveys and in-person studies, due to underlying specific methods for participation that might create self-selection biases (Brüggen et al., 2011; Matthijsse et al., 2015; Yeh & Hirsch, 2023). Thus, we also included a variable differentiating between online and non-online studies.

Finally, journal quartiles were included to control for potential publication bias, as a potential measure of the quality of scientific peer-reviewed and publication process and, thus, of the scientific quality of the article.

The correlation matrix and variance inflation factor (VIF) led to exclude some specific variables due to multicollinearity issues (e.g., the geographic region was correlated with share of food expenditure in the country).

4. The econometric modelling of the MRA: the PET-PEESE estimator

MRA approximations for publication bias were developed by Stanley and Doucouliagos (2014b), by specifically combining quadratic approximation with the Egger et al. (1997) regression, accommodating also heterogeneity and allowing to distinguish between the funnel asymmetry and publication bias. The Egger linear approximation of reported effects and their standard errors (SE, usually referred to as precision) is used as a test for publication bias:

$$\%WTP_i = \beta_0 + \alpha_0 SE_i + \varepsilon_i \quad (1)$$

where testing for $H_0 : \beta_0 = 0$ is the so-called precision-effect test (PET): whether a genuine empirical effect beyond publication bias exists. As $SE \rightarrow 0$, studies are objectively better, hence the estimated effects should approach, on average, β_0. On the other hand, $H_0 : \alpha_0 = 0$ is a test for publication bias effects, usually referred to as funnel-asymmetry test (FAT)[4]: when the null is rejected there is evidence of a linear relationship between studies' estimates and SE, hence publication bias (Costa-Font et al., 2011; Egger et al., 1997; Stanley, 2008, 2017; Stanley & Doucouliagos, 2014b). Therefore, regressing percentage premiums on WTP estimates' standard errors indicates whether a significant and positive relation exists: studies need to search for larger effects to demonstrate statistical significance when standard errors are large (Lin, 2023; Penn & Hu, 2021). As per the relevant literature, we refer to (1) as the FAT-PET specification.

[4] The funnel plot has been often used for detecting publication bias, visualizing the relationship between precision (i.e. $1/SE$) and non-standardized estimate, with studies with less precision (i.e. larger SE) at the bottom of the graph, producing estimates that are more spread out, and the inverted funnel shape relating to the absence of publication bias. Nevertheless, visual inspections can be ambiguous, subjective and deceiving, and formal statistical test of such asymmetry are preferred (Stanley, 2008).

Nevertheless, Stanley (2008) proved that such a linear approximation is a downward-biased estimate of the true effect when this is actually different from zero. Therefore (Stanley & Doucouliagos, 2014b), the proposed approach is the use of higher-order polynomials to better approximate the relationships between WTP estimates and precision, capturing potential non-linearities, replacing SE_i in (1) with the variance of the effect size, SE_i^2:

$$\%WTP_i = \gamma_0 + \alpha_1 SE_i^2 + \varepsilon_i \quad (2)$$

In this specification, α_1 is again a test for publication bias ($H_0 : \alpha_1 = 0$) (Yeh & Hirsch, 2023), while γ_0 is called the precision-effect estimate with SE (PEESE), and it is superior to the PET estimator whenever there exists a genuine non-zero effect (Stanley & Doucouliagos, 2014b). suggest using the combination of the two, the PET-PEESE (conditional) estimator: when PET in (1) rejects the null ($H_0 : \beta_0 = 0$), then the PEESE estimator (2) should be used for a correct estimate of its true effect (i.e. $\widehat{\gamma_0}$). On the contrary, when PET cannot reject the null, i.e. there is no evidence of a nonzero effect, $\widehat{\beta_0}$ is a superior estimator.

Both the linear and quadratic approximations have to be estimated via WLS to control for heteroscedasticity, commonly by $1/\ SE_i^2$ (also referred to as precision-squared) as weights. The large variation among standard errors of reported effects is often observed in MRA, hence defining large heteroscedasticity among reported estimates, which would bias the OLS estimates as the error terms ε_i would not be iid. $1/\ SE_i^2$ gives relatively more (less) weight to the most (least) precise estimates. Therefore, both systematic and random heterogeneity are efficiently handled by the PET-PEESE estimator, leading to a more complex and multiple model including several variables that may explain results variability, from socioeconomic traits to year dummies (Stanley, 2017; Stanley & Doucouliagos, 2015).[5] Accordingly, we will be estimating the following two equations by means of the WLS estimator[6]:

$$\%WTP_i = \beta_0 + \alpha_0 SE_i + \sum_{k=1}^{K} \beta_k X_{ik} + \varepsilon_i \quad (3)$$

$$\%WTP_i = \gamma_0 + \alpha_1 SE_i^2 + \sum_{k=1}^{K} \beta_k X_{ik} + \varepsilon_i \quad (4)$$

where X_{ik} are the independent variables to account for the drivers affecting the %WTP estimates in each study, and β_k is the coefficient for each k characteristics. Overall, especially when the number of studies is large, scholars advocate for running the PET-PEESE estimator for an accurate MRA, and depending on the PET result, using either $\widehat{\beta_0}$ or $\widehat{\gamma_0}$, depending on whether a genuine empirical effect is preset or not, respectively. Finally, even in cases where the publication bias test fails to reject the null (i.e. accepting $\alpha_0 = 0$), because of the low power of the Egger test, $\widehat{\beta_0}$ or $\widehat{\gamma_0}$ should be used (Stanley & Doucouliagos, 2014b). While scholars in the agricultural economics field often used pooled-OLS and RE model (Colen et al., 2018; Lin, 2023; Piracci et al., 2024; Trong Ho et al., 2022), PET-PEESE estimator has been widely proven to provide superior estimates to the widely applied conventional fixed and random effects MRA, even when there are no publication bias effects (Stanley, 2017; Stanley & Doucouliagos, 2014b, 2015). Since one can never rule out the potential presence of publication bias (recalling that tests for such bias have lower power) (Egger et al., 1997; Stanley, 2008), PET-PEESE should always be preferred then RE[7] (although the latter would be the preferred one when no publication bias is present). Simulations proved that even in cases where heterogeneity is high, the PET-PEESE estimator is superior to both RE and FE models (Stanley, 2017).

5. Results

5.1. Descriptive statistics

The descriptive statistics of the included variables in the study are reported in Table 2. The study included 173 papers with a total number of 1065 observations. The relative WTP is, on average, positive, suggesting consumers are willing to pay a relative premium price of 32 % for labeled products. Nevertheless, such value largely varies within a range of −342 % and 624 % across different studies and labels. Indeed, on average the larger relative WTPs are found for Animal Welfare (59 %) and Organic agriculture (42 %) labels, while the lower ones relate to Carbon (12 %) and Health labels (20 %).

We distinguish between 4 categories of explanatory variables: label categories, product characteristics, methodological variables and socioeconomic characteristics. The larger proportion of labels refer to geographical indications, organic agriculture, eco-label and carbon label.

Regarding product categories, fish is the one accounting for a larger number of observations, appearing in 21 % of observations. Vegetables, meat, and cereals follow with observations in 15 %, 12 %, and 10 % of studies. The categories appearing less in the selected papers are “Wine, Beer, and Alcohol” and “Other Food” at 8 % and 4 %, respectively.

We also control for time-related effect, observing that a larger part of the studies were published after 2017, showing that interest in WTP for labeled products in the literature has grown in recent years. Text labels and image labels are split, with 56 % of studies using text descriptions, indicating a reliance on written descriptions of the labels. The Discrete Choice Experiment (DCE) is the dominant method, utilized in 76 % of studies, compared to the less common Contingent Valuation Method (CVM) at 7 % and a remaining category of other methodologies. Online surveys dominate with 74 % of the observations. Regarding journal quartiles, over half of the studies (51 %) were published in Q1 journals, which might suggest a higher quality or impact level. Lastly, a group of socioeconomic and methodological explanatory variables were considered as continuous variables (Table 3). The considered studies included, on average, 2.93 label-specific estimates, reporting 2.5 values of the same label per study. Regarding the socio-economic characteristics, population descriptives suggest that, on average, studies included 53 % of females, with 50 % of the participants with high education, and an average age of 41 years. In addition, the products included had a value of 8.8 euros on average.

The regions of study had a share of country food expenditure on the total GDP between 7 % and 59 %. The study includes observations from all continents, where Europe had the larger total number of observations across the continents, followed by North America and Asia, while Oceania, South America and Africa have fewer observations (see Fig. 2).

5.2. Meta-regression results

Due to the lack of SE reporting for many WTP estimates, we define precision in (1) and (3), as well as precision-squared in (2) and (4), as the square root of the study's observed sample size, $\sqrt{N_i}$ and $\left(\sqrt{N_i}\right)^2$,

[5] As also raised by (Stanley, 2008), within MRA random misspecification and sampling errors are averaged and thereby lessened.

[6] Note that dividing equations (1)–(4) by SE_i and estimated via OLS brings the same results, where the intercept and slope coefficients are reversed. For illustrative purpose, equation (3) would be: $\frac{\%WTP_i}{SE_i} = t_i = \alpha_0 + \beta_0 \frac{1}{SE_i} + \sum_{k=1}^{K} \frac{\beta_k X_{ik}}{SE_i} + \epsilon_i$, while (4) $t_i = \gamma_0 \frac{1}{SE_i} + \alpha_1 SE_i + \sum_{k=1}^{K} \frac{\beta_k X_{ik}}{SE_i} + \epsilon_i$. As in (3) and (4), $H_0 : \alpha_0 = 0$ is FAT, $H_0 : \beta_0 = 0$ is PET and $\widehat{\gamma_0}$ is PEESE (Costa-Font et al., 2011; Stanley, 2017; Stanley & Doucouliagos, 2014a, 2015).

[7] Due to the different weighting of the estimates, as WLS gives more (less) weight to more (less) precise estimates then RE does (Stanley & Doucouliagos, 2014a).

Table 2
Descriptive analysis of the WTP for the meta-regression discrete variables.

	Variables	Description	No. of studies	Number of observations	Share	% WTP (average)	(S. D)	Min	Max
Dependent variable	% WTP	WTP to reference price, percentage	173	1065	1	0.32	0.61	−3.42	6.24
Discrete Independent variables									
Label Categories	Sustainability label[a]	Yes = 1, No = 0	28	77	0.07	0.25	0.63	−1.98	2.94
	Carbon Label	Yes = 1, No = 0	33	112	0.11	0.12	0.28	−0.49	1.96
	Eco-Label	Yes = 1, No = 0	60	159	0.15	0.35	0.61	−0.63	4.14
	Animal Welfare Label	Yes = 1, No = 0	32	70	0.07	0.59	1.13	−2.03	6.24
	Fair Trade Label	Yes = 1, No = 0	46	106	0.10	0.33	0.27	−0.13	1.32
	Geographic Label	Yes = 1, No = 0	56	197	0.18	0.35	0.65	−3.42	5.19
	Health Label	Yes = 1, No = 0	32	128	0.12	0.20	0.32	−0.29	2.16
	Organic Agriculture Label	Yes = 1, No = 0	78	167	0.16	0.42	0.46	−0.47	2.25
	Other label	Yes = 1, No = 0	19	49	0.05	0.25	1.03	−1.43	4.77
Product categories	Vegetables[a]	Yes = 1, No = 0	28	157	0.15	0.28	0.54	−0.63	4.77
	Meat	Yes = 1, No = 0	23	129	0.12	0.27	0.65	−0.95	5.19
	Milk, Honey, Eggs	Yes = 1, No = 0	28	105	0.10	0.40	0.56	−2.03	2.32
	Fish	Yes = 1, No = 0	23	222	0.21	0.19	0.53	−1.98	1.94
	Wine, Beer, Alcohol	Yes = 1, No = 0	22	87	0.08	0.27	0.31	−0.30	1.96
	Coffee and Tea	Yes = 1, No = 0	22	99	0.09	0.46	0.82	−1.22	6.24
	Cereals	Yes = 1, No = 0	22	107	0.10	0.48	0.86	−3.42	4.98
	Fruits	Yes = 1, No = 0	20	117	0.11	0.33	0.33	−0.13	1.60
	Other Food	Yes = 1, No = 0	15	42	0.04	0.40	0.69	−0.06	4.00
Methodological variables	After 2017[a]	Yes = 1, No = 0	142	912	0.86	0.34	0.62	−3.42	6.24
	Before 2017	Yes = 1, No = 0	31	153	0.14	0.22	0.48	−2.03	4.00
	Image or combination text&image labels[a]	Yes = 1, No = 0	69	472	0.44	0.31	0.62	−2.03	6.24
	Text Label	Yes = 1, No = 0	104	593	0.56	0.33	0.59	−3.42	5.19
	Method CE[a]	Yes = 1, No = 0	111	807	0.76	0.31	0.57	−1.98	5.19
	Method CV	Yes = 1, No = 0	24	74	0.07	0.33	0.39	−0.07	1.83
	Other Method	Yes = 1, No = 0	41	184	0.17	0.33	0.80	−3.42	6.24
	Online survey[a]	Yes = 1, No = 0	106	784	0.74	0.29	0.64	−3.42	6.24
	Non-Online Survey	Yes = 1, No = 0	67	281	0.26	0.38	0.48	−0.46	4.00
	Journal Quartile								
	Q1[a]	Yes = 1, No = 0	87	544	0.51	0.27	0.60	−3.42	5.19
	Q2	Yes = 1, No = 0	68	416	0.39	0.40	0.64	−2.03	6.24
	Q3, Q4, not indexed	Yes = 1, No = 0	18	105	0.10	0.24	0.40	−1.22	2.04

[a] Denotes baseline categories.

Table 3
Descriptive statistics of the continuous variables included in the meta-regression analysis.

	Description	No. of studies	No. of observations	Average	Standard deviation	Min	Max
Continuous Independent variables							
Methodological variables	Number of Different Labels evaluated in the study	173	1065	2.93	1.29	1.00	6.00
	Number of reported measures with the Same Label	173	1065	2.51	1.57	1.00	10.00
	Sample size	173	1065	725.99	1455	33	22539
Socioeconomic characteristics of the sample	Females (%)[a]	164	991	0.53	0.09	0.02	0.91
	High Education Level (%)[a]	134	751	0.50	0.15	0.01	1.00
	Mean Age (%)[a]	155	951	41.56	6.47	20.00	68.39
	Country food expenditure	173	1065	0.13	0.06	0.07	0.59
	Reference price product (adjusted PPP)	173	1065	8.87	8.27	0.27	56.82

[a] For papers with missing socioeconomic information, missing values were imputed with the mean of the observed values of the sample.

respectively, as widely applied in the literature (Lin, 2023; Penn & Hu, 2021; Stanley, 2005; Trong Ho et al., 2022; Yeh & Hirsch, 2023). Likewise, as in (Penn & Hu, 2021; Yeh & Hirsch, 2023), we used the inverse root of the sample size as the weights in the WLS estimator for both FAT-PET and PET-PEESE models. Employing sample size directly may correlate positively with publication bias, leading all other variables to correlate with it when used as the weight, hence losing efficiency (Penn & Hu, 2021). As first observed by (Begg & Berlin, 1988), publication bias is proportional to the standard error, as it is proportional to the inverse of the square root of the sample size (in turn related to the inverse of the SE (Stanley, 2005).[8]

Furthermore, since we rely on multiple estimates from the same study (on average, papers report 2.93 estimates), errors in both (3) and (4) may not be independent; hence, clustering at the study level is applied to account for within-study dependence (Colen et al., 2018; Oczkowski & Doucouliagos, 2015; Penn & Hu, 2021; Piracci et al., 2024;

[8] Stanley (2005) shows that the correlation between the inverse of the standard error and the square root of sample size is 0.9.

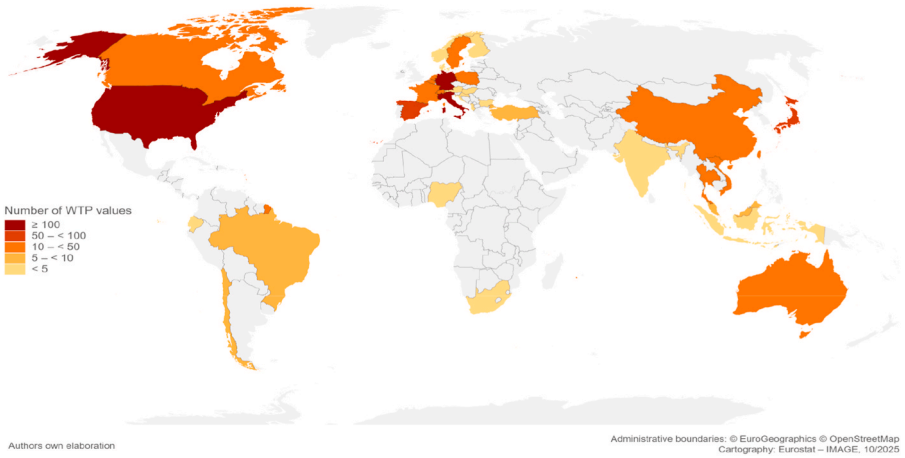


Fig. 2. Overview of the distribution of values of WTP included in the study.
Note: This map indicates the distribution of the measures in the country of origin. The largest number of observations is found in Europe (58 %), North America (23 %) and Asia (14 %). The rest of observations are found in Oceania (4 %), South America (1 %), and Africa (1 %).

Santeramo & Lamonaca, 2019, 2020; Yeh & Hirsch, 2023).[9]

The simplified results are presented in Table 4, displaying the coefficient, standard errors (in brackets) and significance level of the included variables. The first two models provide the results of the FAT-PET (3) and the PET-PEESE (4) to assess the presence of publication bias. The null $H_0 : \alpha_1 = 0$ cannot be rejected, implying that there is no evidence for the presence of publication bias, in line with the most recent similar MRA of Piracci et al. (2024) [10]. The significant constant coefficient in the FAT-PET and the PET-PEESE model confirms the presence of a significant relative WTP for sustainable labels after controlling for publication bias. Both constant terms suggest that consumers are willing to pay a 29 % price premium for products with a sustainable label, which is similar to the 32 % average of the considered sample.

Regarding the model diagnostics, it must be noted that for both models, the R-squared suggests that less than 0.1 % of the variability in relative WTP for sustainable labels can be explained with these simple models. Models 1 to 4 take into account the underlying characteristics of the considered papers, allowing for a deeper understanding of which factors might explain the variation in the relative WTP for sustainable labels.

The results of the full MRA indicate label categories, food categories, reference price, the country-specific food expenditure, high education, studies based on CVMs and papers published before 2017 as sources of heterogeneity, hence exerting a significant effect.

Results suggest that there are significant differences when comparing the percentage price premium of the generic Sustainability label to other more specific label categories, namely Carbon, Animal Welfare and Geographic claims. Geographical and Animal Welfare labels score significantly positively whereas, interestingly, the MRA does not deliver any significant coefficient for Organic nor Eco-labels, hence suggesting that consumers do not bear a different WTP when comparing the more generic sustainability claim with these two specific labels.

Significant preferences are found related to food product categories. When compared to vegetable products, a significant and positive relative WTP is found for labels placed on all remaining product categories, but fruits and other products.

Results also suggest that a lower %WTP is found for products with higher prices (see variable Reference price product (adjusted PPP)). Other significant differences are found for high education (negative effect), country food expenditure, use of CVMs and papers published before 2017.

All models showed p-values (Prob > F) below 0.005, suggesting a relationship between the selected independent variables and the percentage price premium of sustainability labels. The coefficient of the precision proxy $\left(1/\sqrt{N_i}\right)^2$ was not found significant in any of the models, once again confirming the absence of publication bias. The R-squared ranged between 0.09 and 0.19, indicating that the suggested regressions explain between 9 % and 19 % suggesting that model 4 with the full specification provides a better overall fit.

6. Discussion

The results of this meta-analysis provide a comprehensive overview of the effects of sustainable food labeling on consumers' WTP. The MRA performed confirms and extends the findings of earlier reviews, providing substantial originality and added value compared to the literature on sustainable food labeling and consumer WTP from different perspectives. We have conducted an up-to-date large-scale review, considering more than 1000 observations stemming from 173 reviewed studies, and applied state-of-the-art econometric models, incorporating a broad range of label types, food categories, socio-demographic and methodological variables, accounting for potential publication bias.

Results highlighted relevant heterogeneity in consumers' preferences, which also challenges previous studies and common assumptions. Consistent with the previous findings (e.g., Piracci et al., 2024), our results show a positive and significant consumer WTP for sustainability food labels, with an average relative premium of around 30 % (29–32 %). This finding aligns with common evidence that consumers value food products with sustainability attributes on labels: for instance, Li and Kallas (2021) found an average WTP 29.5 % for sustainable food attributes, Piracci et al. (2024) indicated an average 27 % premium for sustainable labels, Yeh and Hirsch (2023) shown that consumers are WTP 26 % more for country-of-origin labels or Yang and Renwick (2019) that found 35.8 % WTP a price premium for Organic label and 31.9 % for Animal Welfare. However, we also evidenced a significant heterogeneity across the WTP estimates, from strongly negative to very high positive premiums. In some cases, such as for Carbon footprint labels, SLs are linked to negative WTPs. It is often the case when such labels show the (eventually higher) carbon footprint of a product compared to its counterpart (e.g., a no-label product). This significantly negative coefficient for Carbon labels also confirms what was previously

[9] As per Angrist and Pischke (2008), at least 42 clusters are needed when using cluster robust standard errors.

[10] Further MRAs on WTP for food products also found no publication bias (Lin, 2023; Oczkowski & Doucouliagos, 2015; Santeramo & Lamonaca, 2019, 2020; Yeh & Hirsch, 2023). Significant publication selection is found in the similar MRA of (Li & Kallas, 2021)

Table 4
Result of Meta-Regression analysis.

	PET-PEESE Models Results											
		fat_peese		pet_peese		model 1		Model 2		Model 3		Model 4	
	Precision proxy	0.001											
		(0.003)											
	Precision Proxy^2			0.000		0.000		0.000		0.000		0.000	
				(0.000)		(0.000)		(0.000)		(0.000)		(0.000)	
	Constant	0.288	***	0.289	***	0.238	***	0.467	*	0.203	**	0.343	
		(0.070)		(0.050)		(0.063)		(0.273)		(0.093)		(0.338)	
Label Categories[a]	Carbon Label					−0.154	*	−0.190	**	−0.156	*	−0.200	**
						(0.080)		(0.095)		(0.082)		(0.098)	
	Eco- labels Label					0.055		0.054		0.031		0.041	
						(0.086)		(0.091)		(0.084)		(0.088)	
	Animal Welfare Label					0.479	**	0.505	**	0.429	*	0.433	*
						(0.229)		(0.237)		(0.232)		(0.242)	
	Fair Trade & labor Label					0.106		0.101		0.034		0.033	
						(0.094)		(0.096)		(0.120)		(0.109)	
	Geographic Label					0.159	*	0.159	*	0.180	**	0.174	*
						(0.083)		(0.087)		(0.086)		(0.095)	
	Health & Nutrition Label					−0.076		−0.038		−0.087		−0.072	
						(0.069)		(0.095)		(0.073)		(0.101)	
	Organic Agriculture					0.156		0.145		0.135		0.128	
						(0.100)		(0.094)		(0.112)		(0.101)	
	Other label					−0.255		−0.265		−0.256		−0.263	
						(0.263)		(0.260)		(0.260)		(0.270)	
Socio economic characteristics	Females (%)							−0.086				−0.084	
								(0.198)				(0.197)	
	High Education Level (%)							−0.320	**			−0.416	***
								(0.153)				(0.139)	
	Mean Age (%)							−0.003				−0.000	
								(0.005)				(0.005)	
	Country Food Expenditure							1.205	*			1.155	*
								(0.637)				(0.620)	
Methodological variables	Published before 2017							−0.155	**			−0.156	**
								(0.069)				(0.063)	
	Text Label							−0.019				−0.034	
								(0.076)				(0.064)	
	Contingent Valuation							−0.190	**			−0.192	**
								(0.087)				(0.081)	
	Other Methods							0.109				0.044	
								(0.106)				(0.100)	
	Number of different labels valued in the study							−0.030				−0.034	
								(0.023)				(0.027)	
	Number of measures with Same Label							0.009				0.017	
								(0.025)				(0.021)	
	non-Online Survey							0.039				0.026	
								(0.079)				(0.084)	
	Journal Quartile[b]												
	2. Q2							0.013				0.032	
								(0.098)				(0.075)	
	3. Q3+Q4+no index							−0.041				−0.055	
								(0.110)				(0.093)	
Product categories[c]	Meat									0.215		0.281	*
										(0.180)		(0.167)	
	Milk, Honey, and Eggs									0.160		0.217	*
										(0.111)		(0.115)	
	Fish									0.074		0.159	
										(0.072)		(0.113)	
	Wine, Beer, and Alcohol									0.138	*	0.206	**
										(0.078)		(0.097)	
	Coffee and Tea									0.185		0.264	*
										(0.134)		(0.155)	
	Cereals									0.249	**	0.334	***
										(0.111)		(0.116)	
	Fruits									0.045		0.162	*
										(0.072)		(0.087)	
	Other Food									0.208		0.214	
										(0.142)		(0.144)	
	Reference price product (adjusted PPP)									−0.009	***	−0.009	***
										(0.003)		(0.003)	
	Number of observations	1065		1065		1065		1065		1065		1065	
	F-Stat	0.22		0.53		3.94		3.91		4.49		4.68	
	Prob > F	0.64		0.47		0.0001		0.0000		0.0000		0.0000	
	R squared	0.0005		0.0009		0.0937		0.1515		0.1382		0.1979	

***p < .01, **p < .05, *p < .1.
[a] Sustainability label as reference.
[b] Published in Q1 Journals as reference.
[c] Vegetables as reference.

found in the literature, where this type of environmental attribute has shown lower relative importance levels compared to others (Apostolidis & McLeay, 2016; Meyerding, 2016; Yang & Renwick, 2019).

On the other hand, Geographical labels score significantly positively. Several studies have already documented the strong role of the labeled domestic origin on consumer choices, which is often perceived as an indicator of the healthfulness and safety of the product (Santeramo et al., 2020). Likewise, the positive and significant WTP for Animal Welfare labels confirms the results of the meta-analysis performed on WTP for farm animal welfare standards by (Lagerkvist & Hess, 2011), and the predicted premium by Yang and Renwick (2019). Thus, these results confirm how ethical and quality perceptions are key elements in driving consumer valuation as mentioned in previous works (Lagerkvist & Hess, 2011; Piracci et al., 2024; Santeramo & Lamonaca, 2020, 2021; Yang & Renwick, 2019). As already discussed, the negative premiums associated with Carbon labels, as also evidenced by Majer et al. (2022), may elicit the contrasting effects of positive-negative eco-labelling schemes when a more favorable (green classes) or unfavorable (red classes) categorization might be associated with, respectively, positive or negative WTP values. In addition, from this perspective, other scholars, such as Apostolidis and McLeay (2016), noted limited consumer WTP specifically for carbon footprint information, possibly due to perceived complexity or negative associations. In contrast with other reviews (Bastounis et al., 2021; Katt & Meixner, 2020; Yang & Renwick, 2019), our findings did not find a significant effect of Organic and Eco-labels. Although a low WTP for environmentally friendly attributes was also reported by another meta-analysis (Li & Kallas, 2021), it is possible that considering these labels along with other different sustainability labels, including ethical and geographical ones, might have covered their effect of eco-friendly features on consumers' WTP. Similarly, no significant effect was observed for Health & Nutrition labels. This finding may reflect the high heterogeneity within this category, including different labelling systems such as Nutri-Score, health or nutrient warning and claims, traffic light labels, etc. Song et al. (2021) in their systematic review, found that Nutri-Score significantly reduced consumers' WTP for less healthful products, whereas no significant effect was observed for textual nutritional or health warnings. Moreover, they reported no evidence of a change in WTP when the more healthful products are labeled with traffic light labels, or health/nutrition claims. Eventually, more specific and detailed analysis could yield significant results. For instance, Dolgopolova and Teuber (2018) reported that the specific health benefit of "lowering cholesterol" was associated with significantly higher WTP values compared to all other health benefits included in their meta-analysis.

The MRA also identifies significant variation in consumers' WTP across food types, with meat, dairy, wine and beer, coffee, and cereals generally commanding higher premiums compared to vegetables. Products like dairy, coffee, wine and cereal often have better consumer understanding and appreciation in the market, leading to consumers being more inclined to pay a premium for labels that assure origin or ethical production. Similarly, consumers may perceive meat, dairy, wine, and coffee as having higher environmental or social impacts compared to vegetables, thus resulting in consumers' willingness to pay for a higher premium for sustainability labels that contribute to address such impacts. Although Piracci et al. (2024) did not retrieve any significant effect of product category on consumers' WTP for SLs, Yeh and Hirsch (2023), for instance, found that WTP is higher when a Geographical label is used on animal-based compared to plant-based products. On the contrary, Li and Kallas (2021) indicated significantly lower WTP estimates for drink products than for dairy and fruit & vegetable products. Since our categorization of food types was more specific compared to the other reviews, the significant variability of WTP by food types may confirm different consumer perceptions of sustainability across various products. However, more specific research is needed to understand if this result might indicate a growing consumer awareness of the variability in the environmental impact among different food categories (Crippa et al., 2021).

Several methodological variables might affect WTP estimates. Consistent with other reviews (Bastounis et al., 2021; Li & Kallas, 2021; Piracci et al., 2024; Yeh & Hirsch, 2023), our results indicate the growing importance of DCEs and online surveys in recent research, reflecting methodological trends that affect WTP estimates. In this context, we found that the use of Contingent valuation methods (CVMs) and the reference product price show significant and negative associations with consumers' WTP. The negative correlation of CVM is somewhat surprising and conflicting with the findings of the meta-analysis performed by Li and Kallas (2021), who showed that the WTP estimate for sustainable food products obtained from the CVM (38.3 %) was higher than that from the DCEs (27.9 %). However, a similar result was already obtained in other domains: for instance, Printezis et al. (2019) indicate that employing DCEs can lead to higher WTP estimates for local food as compared to the application of other experimental techniques, including CVMs. Similarly (Guo et al., 2023), found that values of marine ecosystem services obtained by the CVM were lower than those acquired through the DCE method. Evidently, the common notion that choice-based elicitation mechanisms are associated with less hypothetical bias (Murphy et al., 2005) is somehow outdated by the present results, indicating that when clearly stated elicitation mechanisms are used, the CVM might provide even lower WTP estimates compared to other methods. Moreover, our meta-regression did not find significant publication bias, contradicting concerns raised in the agricultural economics debate (Ferraro & Shukla, 2023; Josephson & Michler, 2018; Rommel & Weltin, 2021), possibly suggesting an improved research transparency or more robust study design methods in the sustainable labeling field. The negative association between product reference price and WTP premium corroborates insights from Havránek et al. (2020), reinforcing the importance of considering relative rather than absolute price premiums in cross-study comparisons. Moreover, our results reinforce the findings of Kilders and Caputo (2024), who suggested that baseline price sensitivity is a critical factor that is generally omitted in WTP studies (Li & Kallas, 2021; Piracci et al., 2024).

Methodological variables also reveal that studies conducted after 2017 report higher WTP values, somehow contradicting other meta-analyses that did not find any significant effect of the date of publishing (e.g., Li & Kallas, 2021; Piracci et al., 2024). This finding, coherent with the results obtained by Yang and Renwick (2019) indicating an increasing trend of WTP a price premium over time for credence attributes, possibly highlights the increased consumer awareness of sustainability issues. This trend aligns with recent evidence showing that consumer concern for environmental sustainability has intensified over the past decade. For example, a 2022 survey on Italian consumers found that over 60 % reported greater awareness of environmental issues and increased green purchases (Dangelico et al., 2022). Similarly, the 2024 Eurobarometer survey showed that 58 % of EU citizens have already changed their consumption behavior to reduce their environmental impact, including preferring sustainable alternatives (European Commission, 2024).

Furthermore, the results of the present meta-analysis indicate that higher education correlates with lower WTP. This finding is somewhat unexpected, although supported by evidence in the literature. For instance, Bastounis et al. (2021) showed that more educated consumers have lower WTP for eco-labeled food products, arguing that they may

critically scrutinize label claims or be more price-sensitive. Our study adds to previous reviews in other food-related domains, which often reported positive or neutral effects of education on WTP, suggesting that the complexity and variability in consumer responses to SLs might be related to demographic factors, including education. Moreover, our results also indicate a positive and significant effect of country food expenditure on consumers' WTP. While direct evidence connecting country food expenditure to WTP for SLs is limited, other reviews suggested that economic factors, such as income and consumer values, significantly influence attitudes and preferences (Cook et al., 2023). Our finding is similar to the one reported by Li and Kallas (2021), indicating that Asian consumers exhibited higher premiums for food products with sustainable labeling than those in North America and similar to those in Europe, which, in general, report lower food expenditures in relative (percentage) terms. Therefore, the present study adds to the broader literature, indicating that countries with higher food expenditure may be associated with increased WTP for sustainability attributes.

In summary, our meta-analysis expands the current literature by systematically addressing less-investigated issues, such as publication bias, employing comprehensive (socio-demographic and methodological) moderator variables, and presenting updated evidence reflecting evolving consumer attitudes and values. While, in general, we provide results consistent with prior findings, our work indicates further specific details, such as the negative impact of Carbon labels and the role of education, that deserve further research to better tailor sustainable labeling policies and marketing strategies.

This study presents several notable strengths. First, it offers a comprehensive and up-to-date meta-analysis of SLs, capturing the growing global interest in sustainable food consumption. Unlike previous reviews that often focused on a single label type or product category (as previously noted by Piracci et al., 2024), this work spans a wide variety of SLs and food products, thereby providing a broader understanding of consumer responses. In addition, the analysis is not limited to DCEs but also incorporates other methodological approaches, offering a more nuanced and realistic picture of how consumers react to SLs in various contexts. This integrative approach allows for a deeper examination of the factors driving consumers' WTP, the role of competing quality cues, and the trade-offs consumers are willing to make. As a result, the findings deliver valuable, policy-relevant insights into the conditions under which SLs are most effective and how they might support the transition to greener consumption patterns.

Despite its contributions, this study is not without limitations. A primary concern lies in the inclusion of both hypothetical and non-hypothetical measures to derive WTP. Hypothetical scenarios, while useful for isolating specific attributes, often simplify the complexity of real-world purchasing environments and may lead to hypothetical bias in consumer responses to sustainability information (Schmidt & Bijmolt, 2020). Moreover, for the WTP estimates, the heterogeneity of the methods (among others, samples' representativeness, econometric models and estimators) might have affected the original estimates. Finally, the scope of the available literature still presents certain constraints. Although the number of studies on SLs has grown in recent years, some areas remain underrepresented, particularly in terms of longitudinal evidence, real-world settings, and low- and middle-income country contexts. As such, further research is needed to expand scientific evidence and capture a more complete picture of WTP heterogeneity across different consumer segments and market environments. Future meta-analyses may benefit from incorporating a broader temporal range and additional moderator variables, as more primary studies become available, in order to strengthen the generalizability and applicability of findings.

7. Conclusions

In response to growing environmental concerns and the global push for sustainable development, this study provides an extensive and up-to-date meta-analysis on the effectiveness of sustainable labeling schemes in shaping consumer food choices. By synthesizing findings across a broad range of labels, food products, and methodological approaches, this work offers a holistic and realistic understanding of consumer choices and intentions on food SLs. The inclusion of multiple study designs and product contexts enables the identification of key drivers of consumers' WTP for sustainability attributes, providing crucial insights for both policymakers and industry stakeholders aiming to design more effective and consumer-aligned labeling strategies.

The meta-regression results confirm a positive consumer response: an average relative WTP of 32 %, closely aligned with the 29 % price premium observed after correcting for publication bias. However, label type, product category, and contextual factors (such as reference prices and food expenditure) do play a critical role in determining consumers' WTP. This heterogeneity underlines the complexity of consumer reactions to sustainability information.

Overall, our findings highlight the substantial influence of SLs on consumer purchasing intentions and its potential influence on the final consumer market. The observed variations underscore the need for tailored communication strategies that consider label type, product category, and consumer characteristics to maximize the impact of sustainability labels on market outcomes. Moreover, results clearly show how social sustainability labels (i.e. animal welfare and geographical origin) are clearly distinguishable from the consumer perspective, indeed entailing a significant (positive) relative WTP when related to generic sustainability claims. Interestingly, this is not entirely true for other environmental-specific labels: while carbon-specific labels are evaluated more negatively by consumers than more holistic and general sustainability claims, organic agriculture and eco-labels do not entail any significant difference in terms of relative WTP when compared with holistic sustainability labels. Overall, this may suggest that a homogeneous sustainable labelling initiative may not work in enhancing the consumers' WTP, as there is significant heterogeneity across sustainability dimensions in terms of WTP for each of them. We find that some stand-alone sustainability aspects carry a higher premium than more encompassing ones, even if some (i.e. Carbon labels) show the opposite. When designing a holistic sustainability label, beyond the technical aspects of how to aggregate impacts, this diversity in consumer preferences adds a new layer of complexity when trying to maximize the market support for such labels.

Finally, this meta-analysis questions the potential of a comprehensive holistic sustainability labeling as a different market-based strategy when compared to other existing labels associated with sustainability aspects (organic production, animal welfare, geographic, health, carbon or eco-labels). According to the results, the implementation of such a comprehensive system will not lead to significantly different WTP towards existing labels, despite the potential simplification for consumers of a less-scattered sustainable labelling system. The implementation of such a label will not only require continuing efforts towards the design of such a comprehensive label providing an overarching assessment of different sustainability dimensions, but it will also entail the need for a communication strategy to differentiate this labeling strategy from other existing labels.

CRediT authorship statement

GA, JTC, FA, JBH, GS, DM conceived the design of the study (conceptualization and methodology) and carried out the research work. GA, JTC, FA, GS, DM analyzed the data and wrote the manuscript. JBH contributed to the data analysis, reviewed and edited the manuscript. All authors contributed to the final edition and approval of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.tifs.2025.105435.

Data availability

Data will be made available on request.

References

Abdu, N., & Mutuku, J. (2021). Willingness to pay for socially responsible products: A meta– analysis of coffee ecolabelling. *Heliyon, 7*(6), Article e07043.

Ammann, J., Arbenz, A., Mack, G., Nemecek, T., & El Benni, N. (2023). A review on policy instruments for sustainable food consumption. *Sustainable Production and Consumption, 36*, 338–353.

Apostolidis, C., & McLeay, F. (2016). Should we stop meating like this? Reducing meat consumption through substitution. *Food Policy, 65*, 74–89.

Arpinon, T., & Lefebvre, M. (2025). Registered reports and associated benefits for agricultural economics. *Q Open, 5*(3), Article qoae011.

Bastounis, A., Buckell, J., Hartmann-Boyce, J., Cook, B., King, S., Potter, C., Bianchi, F., Rayner, M., & Jebb, S. A. (2021). The impact of environmental sustainability labels on willingness-to-pay for foods: A systematic review and meta-analysis of discrete choice experiments. *Nutrients, 13*(8), 2677.

Begg, C. B., & Berlin, J. A. (1988). Publication bias: A problem in interpreting medical data. *Journal of the Royal Statistical Society: Series A, 151*(3), 419. https://doi.org/10.2307/2982993

Bødker, M., Pisinger, C., Toft, U., & Jørgensen, T. (2015). The rise and fall of the world's first fat tax. *Health Policy, 119*(6), 737–742. https://doi.org/10.1016/j.healthpol.2015.03.003

Brüggen, E., Wetzels, M., De Ruyter, K., & Schillewaert, N. (2011). Individual differences in motivation to participate in online panels: The effect on reponse rate and reponse quality perceptions. *International Journal of Market Research, 53*(3), 369–390.

Caussade, S., Ortúzar, J. de D., Rizzi, L. I., & Hensher, D. A. (2005). Assessing the influence of design dimensions on stated choice experiment estimates. *Transportation Research Part B: Methodological, 39*(7), 621–640. https://doi.org/10.1016/j.trb.2004.07.006

Christensen, G., & Miguel, E. (2018). Transparency, reproducibility, and the credibility of economics research. *Journal of Economic Literature, 56*(3), 920–980.

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

Colen, L., Melo, P. C., Abdul-Salam, Y., Roberts, D., Mary, S., & Gomez Y Paloma, S. (2018). Income elasticities for food, calories and nutrients across Africa: A meta-analysis. *Food Policy, 77*, 116–132. https://doi.org/10.1016/j.foodpol.2018.04.002

Cook, B., Costa Leite, J., Rayner, M., Stoffel, S., Van Rijn, E., & Wollgast, J. (2023). Consumer interaction with sustainability labelling on food products: A narrative literature review. *Nutrients, 15*(17), 3837.

Costa-Font, J., Gemmill, M., & Rubert, G. (2011). Biases in the healthcare luxury good hypothesis?: A meta-regression analysis. *Journal of the Royal Statistical Society - Series A: Statistics in Society, 174*(1), 95–107. https://doi.org/10.1111/j.1467-985X.2010.00653.x

Covidence systematic review. (2020). *Covidence systematic review*. Melbourne, Australia: Veritas Health Innovation. . (Accessed 17 March 2020).

Crippa, M., Solazzo, E., Guizzardi, D., Monforti-Ferrario, F., Tubiello, F. N., & Leip, A. (2021). Food systems are responsible for a third of global anthropogenic GHG emissions. *Nature Food, 2*(3), 198–209.

Dangelico, R. M., Schiaroli, V., & Fraccascia, L. (2022). Is Covid-19 changing sustainable consumer behavior? A survey of Italian consumers. *Sustainable Development, 30*(6), 1477–1496.

Dendler, L. (2014). Sustainability Meta labelling: An effective measure to facilitate more sustainable consumption and production? *Journal of Cleaner Production, 63*, 74–83.

Deselnicu, O. C., Costanigro, M., Souza-Monteiro, D. M., & McFadden, D. T. (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.

Dolgopolova, I., & Teuber, R. (2018). Consumers' willingness to pay for health benefits in food products: A meta-analysis. *Applied Economic Perspectives and Policy, 40*(2), 333–352.

Egger, M., Smith, G. D., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ, 315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629

European Commission. (2020). *Farm to fork strategy for a fair, healthy and environmentally-friendly food system*. Publications Office of the European Union. https://ec.europa.eu/food/system/files/2020-05/f2f_action-plan_2020_comm2020_en.pdf.

European Commission. (2024). Attitudes of Europeans towards the environment. *Eurobaromenter Report, Special Eurobarometer, 550*. https://doi.org/10.2779/07854

FAO, F. and A. O. of the U. N. (2018). *Sustainable food systems: Concept and framework*. Rome: FAO.

Ferguson, J., Littman, R., Christensen, G., Paluck, E. L., Swanson, N., Wang, Z., Miguel, E., Birke, D., & Pezzuto, J.-H. (2023). Survey of open science practices and attitudes in the social sciences. *Nature Communications, 14*(1), 5401.

Ferraro, P. J., & Shukla, P. (2023). Credibility crisis in agricultural economics. *Applied Economic Perspectives and Policy, 45*(3), 1275–1291. https://doi.org/10.1002/aepp.13323

Finger, R., Henningsen, A., Höhler, J., Huber, R., Rommel, J., & Grebitus, C. (2025). Open science in agricultural economics. *Q Open, 5*(3), Article qoae029.

Guo, J., Chen, K., & Hynes, S. (2023). Multilevel meta-analysis of stated preference values of marine ecosystem services in China. *Marine Development, 1*(1), 8.

Havránek, T., Stanley, T. D., Doucouliagos, H., Bom, P., Geyer-Klingeberg, J., Iwasaki, I., Reed, W. R., Rost, K., & van Aert, R. C. M. (2020). Reporting guidelines for meta-analysis in economics. *Journal of Economic Surveys, 34*(3), 469–475. https://doi.org/10.1111/joes.12363

Josephson, A., & Michler, J. D. (2018). Viewpoint: Beasts of the field? Ethics in agricultural and applied economics. *Food Policy, 79*, 1–11. https://doi.org/10.1016/j.foodpol.2018.08.001

Katt, F., & Meixner, O. (2020). A systematic review of drivers influencing consumer willingness to pay for organic food. *Trends in Food Science & Technology, 100*, 374–388. https://doi.org/10.1016/j.tifs.2020.04.029

Kilders, V., & Caputo, V. (2024). A reference-price-informed experiment to assess consumer demand for beef with a reduced carbon footprint. *American Journal of Agricultural Economics, 106*(1), 3–20.

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

Li, S., & Kallas, Z. (2021). Meta-analysis of consumers' willingness to pay for sustainable food products. *Appetite, 163*, Article 105239. https://doi.org/10.1016/j.appet.2021.105239

Lin, W. (2023). The effect of product quantity on willingness to pay: A meta-regression analysis of beef valuation studies. *Agribusiness, 39*(3), 646–663. https://doi.org/10.1002/agr.21798

Lusk, J. L., Jamal, M., Kurlander, L., Roucan, M., & Taulman, L. (2005). A meta-analysis of genetically modified food valuation studies. *Journal of Agricultural and Resource Economics, 30*(1), 28–44.

Majer, J. M., Henscher, H. A., Reuber, P., Fischer-Kreer, D., & Fischer, D. (2022). The effects of visual sustainability labels on consumer perception and behavior: A systematic review of the empirical literature. *Sustainable Production and Consumption, 33*, 1–14.

Matthijsse, S. M., De Leeuw, E. D., & Hox, J. J. (2015). Internet panels, professional respondents, and data quality. *Methodology, 11*(3), 81–88. https://doi.org/10.1027/1614-2241/a000094

Meyerding, S. G. H. (2016). Consumer preferences for food labels on tomatoes in Germany – A comparison of a quasi-experiment and two stated preference approaches. *Appetite, 103*, 105–112. https://doi.org/10.1016/j.appet.2016.03.025

Murphy, J. J., Allen, P. G., Stevens, T. H., & Weatherhead, D. (2005). A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics, 30*, 313–325.

Oczkowski, E., & Doucouliagos, H. (2015). Wine prices and quality ratings: A meta-regression analysis. *American Journal of Agricultural Economics, 97*(1), 103–121. https://doi.org/10.1093/ajae/aau057

Penn, J. M., & Hu, W. (2021). The extent of hypothetical bias in willingness to accept. *American Journal of Agricultural Economics, 103*(1), 126–141. https://doi.org/10.1111/ajae.12121

Piracci, G., Lamonaca, E., Santeramo, F. G., Boncinelli, F., & Casini, L. (2024). On the willingness to pay for food sustainability labelling: A meta-analysis. *Agricultural Economics, 55*(2), 329–345. https://doi.org/10.1111/agec.12826

Potter, C., Bastounis, A., Hartmann-Boyce, J., Stewart, C., Frie, K., Tudor, K., Bianchi, F., Cartwright, E., Cook, B., Rayner, M., & Jebb, S. A. (2021). The effects of environmental sustainability labels on selection, purchase, and consumption of food and drink products: A systematic review. *Environment and Behavior, 53*(8), 891–925. https://doi.org/10.1177/0013916521995473

Printezis, I., Grebitus, C., & Hirsch, S. (2019). The price is right!? A meta-regression analysis on willingness to pay for local food. *PLoS One, 14*(5), Article e0215847. https://doi.org/10.1371/journal.pone.0215847

Rommel, J., & Weltin, M. (2021). Is there a cult of statistical significance in agricultural economics? *Applied Economic Perspectives and Policy, 43*(3), 1176–1191. https://doi.org/10.1002/aepp.13050

Rondoni, A., & Grasso, S. (2021). Consumers behaviour towards carbon footprint labels on food: A review of the literature and discussion of industry implications. *Journal of Cleaner Production, 301*, Article 127031.

Ryan, M., & Wordsworth, S. (2000). Sensitivity of willingness to pay estimates to the level of attributes in discrete choice experiments. *Scottish Journal of Political Economy, 47*(5), 504–524.

Santeramo, F. G., & Lamonaca, E. (2019). The effects of non-tariff measures on agri-food trade: A review and meta-analysis of empirical evidence. *Journal of Agricultural Economics, 70*(3), 595–617. https://doi.org/10.1111/1477-9552.12316

Santeramo, F. G., & Lamonaca, E. (2020). Evaluation of geographical label in consumers' decision-making process: A systematic review and meta-analysis. *Food Research International, 131*, Article 108995. https://doi.org/10.1016/j.foodres.2020.108995

Santeramo, F. G., & Lamonaca, E. (2021). Objective risk and subjective risk: The role of information in food supply chains. *Food Research International, 139*, Article 109962. https://doi.org/10.1016/j.foodres.2020.109962

Santeramo, F. G., Lamonaca, E., Carlucci, D., De Devitiis, B., Seccia, A., Viscecchia, R., & Nardone, G. (2020). On the relevance of the region-of-origin in consumers studies. *Bio-based and Applied Economics, 9*(2), 137–154.

Schmidt, J., & Bijmolt, T. H. A. (2020). Accurately measuring willingness to pay for consumer goods: A meta-analysis of the hypothetical bias. *Journal of the Academy of Marketing Science, 48*, 499–518.

Song, J., Brown, M. K., Tan, M., MacGregor, G. A., Webster, J., Campbell, N. R. C., Trieu, K., Mhurchu, C. N., Cobb, L. K., & He, F. J. (2021). Impact of color-coded and warning nutrition labelling schemes: A systematic review and network meta-analysis. *PLoS Medicine, 18*(10), Article e1003765.

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.

Stanley, T. (2001). Wheat from chaff: Meta-analysis as quantitative literature review. *The Journal of Economic Perspectives, 15*(3), 131–150. https://doi.org/10.1257/jep.15.3.131

Stanley, T. D. (2005). Beyond publication bias. *Journal of Economic Surveys, 19*(3), 309–345. https://doi.org/10.1111/j.0950-0804.2005.00250.x

Stanley, T. D. (2008). Meta-regression methods for detecting and estimating empirical effects in the presence of publication selection. *Oxford Bulletin of Economics & Statistics, 70*(1), 103–127. https://doi.org/10.1111/j.1468-0084.2007.00487.x

Stanley, T. D. (2017). Limitations of PET-PEESE and other meta-analysis methods. *Social Psychological and Personality Science, 8*(5), 581–591. https://doi.org/10.1177/1948550617693062

Stanley, T. D., & Doucouliagos, H. (2012). In T. D. Stanley, & H. Doucouliagos (Eds.), *Meta-regression analysis in economics and business* (1st ed.). Routledge.

Stanley, T. D., & Doucouliagos, H. (2014a). Better than random: Weighted least squares meta-regression analysis. *Economic Series*. Economics Working Paper Series SWP 2013/2.

Stanley, T. D., & Doucouliagos, H. (2014b). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5*(1), 60–78. https://doi.org/10.1002/jrsm.1095

Stanley, T. D., & Doucouliagos, H. (2015). Neither fixed nor random: Weighted least squares meta-analysis. *Statistics in Medicine, 34*(13), 2116–2127. https://doi.org/10.1002/sim.6481

Thøgersen, J. (2023). How does origin labelling on food packaging influence consumer product evaluation and choices? A systematic literature review. *Food Policy, 119*, Article 102503.

Tirion, A. S. C., D'Lima, D., Terlet, J., Rao, R., & Gutman, L. M. (2025). Identifying behaviour change techniques for sustainable food consumption: A systematic review using the BCTTv1. *Appetite, 214*, Article 108057. https://doi.org/10.1016/j.appet.2025.108057

Torma, G., & Thøgersen, J. (2021). A systematic literature review on meta sustainability labeling – What do we (not) know? *Journal of Cleaner Production, 293*, Article 126194. https://doi.org/10.1016/j.jclepro.2021.126194

Trong Ho, P., Burton, M., Ma, C., & Hailu, A. (2022). Quantifying heterogeneity, heteroscedasticity and publication bias effects on technical efficiency estimates of rice farming: A meta-regression analysis. *Journal of Agricultural Economics, 73*(2), 580–597. https://doi.org/10.1111/1477-9552.12468

Tully, S. M., & Winer, R. S. (2014). The role of the beneficiary in willingness to pay for socially responsible products: A meta-analysis. *Journal of Retailing, 90*(2), 255–274. https://doi.org/10.1016/j.jretai.2014.03.004

UN. (2015). *Transforming our world: The 2030 agenda for sustainable development*.

Yang, W., & Renwick, A. (2019). Consumer willingness to pay price premiums for credence attributes of livestock products–A meta-analysis. *Journal of Agricultural Economics, 70*(3), 618–639.

Yeh, C., & Hirsch, S. (2023). A meta-regression analysis on the willingness-to-pay for country-of-origin labelling. *Journal of Agricultural Economics, 74*(3), 719–743. https://doi.org/10.1111/1477-9552.12528

Yenipazarli, A. (2015). The economics of eco-labeling: Standards, costs and prices. *International Journal of Production Economics, 170*, 275–286.